

DEVELOPMENT OF A HYBRID QUANTUM MECHANICS-MOLECULAR
MECHANICS MODEL FOR PREDICTING SOLVENT EFFECTS

By

GREG M. PEARL

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1998

ACKNOWLEDGMENTS

First, I would like to acknowledge the person responsible for introducing me to computational chemistry, Dr. Kenny B. Lipkowitz. I would also like to acknowledge Dr. Michael C. Zerner, for his guidance and insight that he provided, along with being a mentor figure. I would like to recognize my unofficial Swedish advisor, Dr. Anders Broo for all of his assistance, scientific conversations, and for the trip to Sweden. In addition I would like to recognize and thank Dr. Erik Deumens, Dr. Daniel Robertson, and Dr. Michael Peterson for assistance and tutelage in matters of chemistry and computers. The Zerner research group should also be acknowledged for their assistance, especially Dr. Marshall Cory and Dan De Kee for editorial comments. I would also like to acknowledge my father, Curtis B. Pearl for his support over the last 26 years. Finally, I would also like to thank the members of my committee; Dr. Monkhorst, Dr. Öhrn, Dr. Richards, Dr. Sabin and Dr. Zerner, for their time and the insight they have provided.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES.....	vi
LIST OF FIGURES.....	viii
ABSTRACT.....	xi
CHAPTER 1: INTRODUCTION.....	1
Overview	1
Molecular Modeling.....	2
Molecular Mechanics.....	2
Quantum Mechanics	3
Modeling Solvent Effects.....	4
Explicit Solvent Modeling.....	5
Electric Field Models.....	7
Hybrid Quantum Mechanics-Molecular Mechanics Methods	8
CHAPTER 2: COMPUTATIONAL MODELS FOR LARGE SYSTEMS.....	11
Introduction	11
Potential Energy Modeling.....	12
Molecular Mechanics.....	12
Quantum Mechanics	15
Semi-Empirical Models	16
Hybrid Quantum Mechanics-Molecular Mechanics Models.....	22
Geometric Modeling Strategies	27
Geometry Optimization	29
Steepest Descent.....	29
Conjugate Gradient.....	30
Quasi-Newton Update Methods.....	31
Molecular Dynamics.....	32
Brownian Dynamics	39
Monte Carlo	40
Boundary Conditions	47
Periodic Boundary Conditions.....	48
Stochastic Boundary Conditions	50
Ewald Summations and Fast Multipole Methods.....	51
Self-Consistent Reaction Fields.....	53

CHAPTER 3: TEST SYSTEMS.....	59
Introduction	59
Molecular Mechanics and Stochastic Methods.....	59
Hybrid Quantum Mechanics-Molecular Mechanics	61
Objective	61
Water Dimer	62
Formamide/Water Complex.....	66
Formamide/Methanol Complex.....	73
Solvent Dimer Testing	78
Water / Ammonia	78
Water/Methanol.....	82
Water/Pyridine	84
Water/Acetone.....	86
Ammonia/Acetone	88
Ammonia/Pyridine	89
Ammonia/Methanol.....	91
Summary and Conclusions.....	93
CHAPTER 4: APPLICATIONS OF THE QM/MM METHOD ON LARGE MOLECULAR SYSTEMS.....	95
Uracil and Dimethyluracil Spectroscopy	95
Background	95
Method	96
Results	97
Summary	108
Solvent Solute Charge Transfer.....	109
Introduction.....	109
Results	110
Summary and Conclusions.....	117
Dissociation Energies.....	118
Introduction.....	118
Results	119
Summary	132
CHAPTER 5: EMPIRICAL BAND WIDTH STUDY	134
Introduction	134
Theory.....	137
Method.....	148
Results.....	150
Benzene and naphthalene in non-polar solvents:.....	150
Comparison of different solvent for pyridine, pyrimidine, and pyridazine:.....	160
Conclusion	181
CHAPTER 6: SUMMARY AND CONCLUSIONS	182

Summary182

Future Work183

APPENDIX A: ZINDO QM/MM User Manual185

APPENDIX B: STOCHASTIC SIMULATION ANALYZER MANUAL ...206

REFERENCES210

BIOGRAPHICAL SKETCH.....225

LIST OF TABLES

Table 1.1: Comparison of Time Required for Molecular Mechanics and Quantum Mechanics Methods.....	6
Table 3.1: Cyclic Water Dimer Results	64
Table 3.2: Summary of the Water Dimer, As Predicted at Different Levels of Theory	65
Table 3.3: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 1.....	70
Table 3.4: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 2.....	71
Table 3.5: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 3.....	71
Table 3.6: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 4.....	72
Table 3.7: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 1.....	75
Table 3.8: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 2.....	76
Table 3.9: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 3.....	76
Table 3.10: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 4.....	77
Table 3.11: Water / Ammonia (W/A) Dimer Results; Conformer 1	80
Table 3.12: Ammonia / Water (A/W) Dimer Results; Conformer 2	81
Table 3.13: Water/Methanol (W/M) Dimer Results; Conformer 1	83
Table 3.14: Methanol/Water (M/W) Dimer Results; Conformer 2	83
Table 3.15: Water/Pyridine (W/P) Dimer Results	85
Table 3.16: Water/Acetone (W/A) Dimer Results.....	87
Table 3.17: Ammonia/Acetone (A/Ace) Dimer Results.....	89
Table 3.18: Ammonia/Pyridine (A/P) Dimer Results.....	90
Table 3.19: Ammonia/Methanol (A/M) Dimer Results; Conformer 1.....	92
Table 3.20: Methanol/Ammonia (M/A) Dimer Results; Conformer 2.....	92
Table 4.1: Comparison of Predicted Absorption Spectra of Uracil with Observed Spectra.....	100
Table 4.2: Comparison of Predicted Absorption Spectra of 1,3-Dimethyluracil with Observed Spectra.....	101
Table 4.3: Calculated Transition Energies for $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex. ..	112
Table 4.4: Comparison of the <i>ab initio</i> and INDO/1 Mulliken Charges for the $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex.....	113
Table 4.5: Calculated QMMM Transition Energies for the $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex.	116
Table 4.6: HF Binding Energy	120

Table 4.7: HCl Binding Energy	121
Table 4.8: HBr Binding Energy.....	122
Table 4.9: H...X Bond Energies.....	126
Table 4.10: Difference in HF Binding Energy When Comparing QM and QM/MM Models.....	128
Table 4.11: Difference in HCl Binding Energy When Comparing QM and QM/MM Models.....	128
Table 4.12: Difference in HBr Binding Energy When Comparing QM and QM/MM Models.....	129
Table 4.13: QM/MM Dissociation Energy of HX Dimers with 4 QM H ₂ O's	131
Table 5.1: Force Constants.....	140
Table 5.2: Calculated Benzene Transitions	155
Table 5.3: Calculated Naphthalene Transitions.....	155
Table 5.4: Comparison of Different Solvents for Pyridine	167
Table 5.5: Comparison of Different Solvents for Pyrimidine	174
Table 5.6: Comparison of Different Solvents for Pyridazine	175
Table A1.1: \$CTRL Keywords	186
Table A1.2: Default Molecular Mechanics Force Field (CFF)	202

LIST OF FIGURES

Figure 2.1: Schematic of Molecular Force Field Expressions.	13
Figure 2.2: One-Electron Two-Center Integral	17
Figure 2.3: One-Electron Two-Center Integral	17
Figure 2.4: Two-Electron Two-Center Integral Differential Atomic and Diatomic Overlap	18
Figure 2.5: Pictorial View of QM/MM Model	23
Figure 2.6: Graphical Representation of how the Various Methods Search Phase Space.	28
Figure 2.7: Steepest Descent Method	30
Figure 2.8: Conjugate Gradient.....	31
Figure 2.9: Types of Monte Carlo Steps.....	44
Figure 2.10: Monte Carlo Time Correlation Plot.....	46
Figure 2.11: Monte Carlo Radial Distribution Function (O...O).....	47
Figure 2.12: 2-D Square Periodic Boundary Conditions.....	49
Figure 2.13: Building up of an Ewald Sum.....	52
Figure 2.14: Self Consistent Reaction Field	54
Figure 2.15: Comparison of the Reaction Field for Ground State and Excited States.....	58
Figure 3.1: Water Dimer C_s Symmetry Conformation	62
Figure 3.2: Water Dimer C_{2v} Symmetry Conformation	63
Figure 3.3: Water Dimer Cyclic Conformation	63
Figure 3.4: Water/Formamide Complex; Conformer 1	66
Figure 3.5: Water/Formamide Complex; Conformer 2	67
Figure 3.6: Water/Formamide Complex; Conformer 3	67
Figure 3.7: Water/Formamide Complex; Conformer 4	68
Figure 3.8: Methanol/Formamide Complex; Conformer 1	73
Figure 3.9: Methanol/Formamide Complex; Conformer 2	73
Figure 3.10: Methanol/Formamide Complex; Conformer 3.....	74
Figure 3.11: Methanol/Formamide Complex; Conformer 4.....	74
Figure 3.12: Water / Ammonia Complex	80
Figure 3.13: Water/Methanol Complex.....	82
Figure 3.14: Water/Pyridine Complex	85
Figure 3.15: Water/Acetone Complex.....	86
Figure 3.16: Ammonia/Acetone Complex.....	88
Figure 3.17: Ammonia/Pyridine Complex	90
Figure 3.18: Ammonia/Methanol Complex.....	91
Figure 4.1: Uracil and 1,3-Dimethyluracil Structures.....	95
Figure 4.2: Hydrogen Bond Distance Between the O7 Carbonyl and the Neighboring Water Hydrogen's as a Function of Simulation Time.....	98
Figure 4.3: Hydrogen Bond Distance Between the O8 Carbonyl and the Neighboring Water Hydrogen's as a Function of Simulation Time.....	99

Figure 4.4: Schematic Drawing of the Keto-Enol Tautomerization.....	105
Figure 4.5: INDO/S-CIS Predicted Spectrum of 1,3-dimethyluracil.....	107
Figure 4.6: Ruthenium (II) Pentaaminopyridine Complex	110
Figure 4.7: Radial Distribution Function for the Amino Groups and Surrounding Water Molecules.....	115
Figure 4.8: QM/MM Predicted Electronic Spectra for $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex.	116
Figure 4.9: HF Potential Energy vs/ Bond Distance.....	123
Figure 4.10: HCl Potential Energy vs/ Bond Distance	124
Figure 4.11: HBr Potential Energy vs/ Bond Distance.....	125
Figure 4.12: Solvent Interactions at Long $\text{H}\cdots\text{X}$ Dissociation Distances...	127
Figure 4.13: Dissociation Energy of HX Dimers + 4 QM H_2O 's.....	130
Figure 5.1: Anatomy of an Ideal Spectral Peak.....	134
Figure 5.2: Band Overlapping	136
Figure 5.3: Effects of the Excited State Surface on Band Width	144
Figure 5.4: Experimental UV Spectra of Benzene	151
Figure 5.5: Calculated UV Spectra of Benzene Using the Band Width Theory	152
Figure 5.6: Calculated UV Spectra of Benzene Using the Band Width Theory and Transition Strength Correction.....	153
Figure 5.7: Calculated UV Spectra of Benzene From Molecular Dynamics Simulation	154
Figure 5.8: Experimental UV Spectra of Naphthalene	156
Figure 5.9: Calculated UV Spectra of Naphthalene Using the Band Width Theory	157
Figure 5.10: Calculated UV Spectra of Naphthalene Using the Band Width Theory and Transition Strength Correction.....	158
Figure 5.11: Calculated UV Spectra of Naphthalene From Molecular Dynamics Simulation.....	159
Figure 5.12: Experimental UV Spectra of Pyridine in Non-polar Solvent .	161
Figure 5.13: Calculated UV Spectra of Gas Phase Pyridine Using the Band Width Theory.....	162
Figure 5.14: Calculated UV Spectra of Pyridine in Non-polar Solvent Using the Band Width Theory	163
Figure 5.15: Experimental UV Spectra of Pyridine in Water	164
Figure 5.16: Calculated UV Spectra of Pyridine Water Using the Band Width Theory.....	165
Figure 5.17: Calculated UV Spectra of Pyridine Water Using the Band Width Theory 2	166
Figure 5.18: Experimental UV Spectra of Pyrimidine in Non-polar Solvent	168
Figure 5.19: Calculated UV Spectra of Gas Phase Pyrimidine Using the Band Width Theory	169
Figure 5.20: Calculated UV Spectra of Pyrimidine in Non-polar Solvent Using the Band Width Theory.....	170
Figure 5.21: Experimental UV Spectra of Pyrimidine in Water	171
Figure 5.22: Calculated UV Spectra of Pyrimidine Water Using the Band Width Theory.....	172

Figure 5.23: Calculated UV Spectra of Pyrimidine Water Using the Band Width Theory 2	173
Figure 5.24: Experimental UV Spectra of Pyridazine in Non-polar Solvent	176
Figure 5.25: Calculated UV Spectra of Gas Phase Pyridazine Using the Band Width Theory	177
Figure 5.26: Calculated UV Spectra of Pyridazine in Non-polar Solvent Using the Band Width Theory.....	178
Figure 5.27: Calculated UV Spectra of Pyridazine Water Using the Band Width Theory.....	179
Figure 5.28: Calculated UV Spectra of Pyridazine Water Using the Band Width Theory 2	180
Figure A1.1: QM/MM Program Flow Chart.....	185
Figure A2.1: Flow Chart for QM/MM Analyzer Program.....	207

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

DEVELOPMENT OF A HYBRID QUANTUM MECHANICS-MOLECULAR
MECHANICS MODEL FOR PREDICTING SOLVENT EFFECTS

By

Greg M. Pearl

May, 1998

Chairman: Dr. Michael C. Zerner
Major Department: Chemistry

Solvent effects can significantly influence the structure, reactivity and spectroscopy of a molecule, and, therefore, proper modeling of these effects is crucial for predicting the behavior of molecules in solution. While this is not the first attempt to incorporate solvent effects into theoretical calculations, the new method presented here has fewer limitations on the types of systems to which it can be applied. This method is based upon calculating the potential energy using a hybrid quantum mechanics-molecular mechanics (QM/MM) methodology for calculating the potential energy. This method models the solvent molecules with molecular mechanics (classical physics) while treating the solute molecule using quantum mechanics (solving the Schrödinger equation). The goal of this hybrid approach is to avoid the "pitfalls" of these two methods while maintaining the advantages of each.

CHAPTER 1 INTRODUCTION

Overview

Chemistry is science that deals with the composition, structure, and properties of substances and with the transformations that they undergo. It is quite broad in that it includes the study of all biological phenomena such as the photophysical properties of photosynthesis [1-5]. In addition, this definition includes any attempts at understanding or predicting the interaction of drugs with living organisms [6,7], synthesis of new compounds such as explosives, plastics, and pigments and dyes [8,9], and for the development of better fertilizers to increase productivity of farm land [10].

Chemists acquire knowledge by utilizing the scientific method; which is the combination of observations, experimentation, and the formulation of laws, hypotheses, and theories. Experimental chemists accomplish this by observing natural phenomena with their instruments. Computational chemists accomplish this by using theoretical models to represent the chemical systems and performing computational simulations using these models.

Computational chemistry is a large field of study, and has been defined as the study of chemistry that involves the use of computers for generation of data, as opposed to using the computer for collecting the experimental results [11]. This field is generally divided into two major

classes: the first class uses databases of acquired chemical knowledge [12] to predict the properties of molecules without concern for the mechanism of how it occurs; the second class simulates chemistry using physical models [13-15] geared toward understanding the physical interpretations or mechanisms of how the phenomena occurs.

This dissertation focuses on the development of computational models for describing the effects of solvation of molecules (Chapter 2, Section 2.4, and Chapter 5). These computation models are then applied to an assortment of “simple” chemical applications for testing the predictive power of these new models (Chapter 3). Finally these new methods are applied to several systems for which the previous methods for including solvent effects have failed to predict experimental observations (Chapter 4).

Molecular Modeling

Molecular modeling is the use of theoretical models for examining the properties of molecules. There are two classes of molecular modeling: molecular mechanics and quantum mechanics. This section will focus upon providing an introduction for both of these types of potential energy calculations.

Molecular Mechanics

Molecular mechanics is an empirical method that uses classical mechanics and electrostatic equations for describing a model system [16]. The use of these classical equations provides a computationally inexpensive method for calculating the potential energy for the system along with analytical gradients of the potential energy with respect to position. Molecular mechanics methods can be used on relatively large systems

because the calculations scale as an N_{MM}^2 procedure, where N_{MM} is the number of interacting particles (atoms or molecules). The total time for a molecular mechanics potential energy calculation is given by

$$t_{MM} = \alpha_{MM} N_{MM}^2 \quad 1.1$$

where α_{MM} is a scaling factor which depends upon the molecular mechanics method. These methods have been successfully applied to several problems such as the explanation of where and how entantioselective binding takes place on permethylated β -cyclodextrin [17]. The molecular mechanics methods usually provide accurate geometries and bulk thermodynamic properties for systems that have been parameterized well, such as the TIPS force field for modeling water [18].

Molecular mechanics methods are limited in the types of problems to which they can be applied. The limitations mostly originate from two sources: the first is the requirement of accurate parameters, the second is a limitation of the model used for describing the system. The second limitation is that electrons are not explicitly modeled and renders molecular mechanics methods inappropriate for modeling systems that involve charge transfer, solvent polarization, electronic spectroscopy, or the formation or dissociation of covalent bonds.

Quantum Mechanics

Quantum mechanics calculations are based upon solution of the Schrödinger Equation [19]. The solution of the Schrödinger Equation is more complicated than the solution of the classical mechanics equations

used for molecular mechanics calculations, and, therefore, these calculations require more time and resources to calculate an equivalent system. The solution of the Schrödinger Equation requires some approximations since an analytic solution is not feasible for molecular systems. The popular methods for approximating the solution to the Schrödinger Equation vary from an N_{QM}^3 to N_{QM}^{10} procedure, where N_{QM} is the number of basis functions (electron orbitals), where the time required for the fastest procedure is given by,

$$t_{QM} = \alpha_{QM} N_{QM}^3 \quad 1.2$$

where α_{QM} is a scaling factor that depends upon the approximate method and program used. The quantum mechanics methods have been successfully applied to several chemical problems [20-24] such as the calculations on cellulose by Green and Zanardi [25].

The limitations of quantum mechanics calculations arise from the time and resources required opposed to limitations imposed by the model. While there are some limitations imposed by the approximate method chosen for solving the Schrödinger Equation, these approximations can be avoided by using a different approximate solution method. The current limitations on quantum mechanics calculations is ~30 or so atoms for *ab initio* calculations or ~1,000 atoms for semi-empirical methods.

Modeling Solvent Effects

It is well known that the environment in which a molecule resides can significantly influence the structure, reactivity, and spectroscopy [26-

30]. It is, therefore, imperative to incorporate these environmental effects into theoretical models when trying to predict the properties of molecules in condensed media (solution). One important type of environmental effect in chemistry is solvation, since most experimental chemistry is performed in condensed media (solvent). Several theoretical models have been developed for predicting solvent effects and these models can be classified into two categories: the first class explicitly models the solvent molecules while the second class uses an electric field to represent the solvent.

Explicit Solvent Modeling

Explicitly modeling the solvent molecules is a logical method for including solvent effects into a theoretical calculation. Although it is very appealing to model every solvent molecule, it is also currently impossible. Experimental chemistry is performed on systems that contain a vast number of particles, while computational chemistry is limited to modeling systems containing much fewer than one million particles (atoms or molecules). Computational chemistry is, therefore, required to use model systems for describing experimental solutions [31-39]. The number of water molecules required to represent a solution varies, although except for all but the simplest models (<15 water molecules) the use of quantum mechanics methods is prohibited. The time factor values in Table 1.1 assume that $\alpha_{QM} = \alpha_{MM}$, which is not correct since $\alpha_{MM} \ll \alpha_{QM}$. The time factor is also comparing the time required for a single potential energy calculation using molecular mechanics to a single self-consistent field (SCF) iteration using quantum mechanics. Several SCF iterations are required for each potential energy calculation using quantum mechanics.

Table 1.1: Comparison of Time Required for Molecular Mechanics and Quantum Mechanics Methods

# waters	Molarity ^A (mol/L)	N_{QM} ^B	N_{MM} ^C	Time Factor ^D
1	5.6×10^1	8	3	5.7×10^1
10	5.6×10^0	80	30	5.7×10^2
56	1.0×10^0	448	168	3.2×10^3
99	5.7×10^{-1}	792	297	5.6×10^3
218	2.6×10^{-1}	1,744	654	1.2×10^4
560	1.0×10^{-1}	4,480	1,680	3.2×10^4
5,600	1.0×10^{-2}	44,800	16,800	3.2×10^5

- A. Molarity assumes a single solute molecule.
 B. Minimal basis set representation using the fastest quantum mechanics method (N_{QM}^3).
 C. Assuming non-rigid model for solvent molecules.
 D. N_{QM}^3 / N_{MM}^2

Modeling the solvent molecules explicitly also introduces the additional theoretical complication of having to search phase space instead of utilizing a geometry optimized structure, due to the large number of local minimum energy conformations present. This is the final “nail in the coffin” for the quantum mechanics methods for explicitly modeling the solvent. Since, the statistical mechanics methods used for searching phase space require ~10,000 potential energy calculation for a 5 ps simulation, while molecular vibrations requires 10-20 ps. If a 20 ps simulation using

molecular mechanics modeling 218 water molecules requires ~1 hour, the same calculation using quantum mechanics would require ~1.5 years.

The explicit modeling of solvent molecules is limited to using molecular mechanics methods and is, therefore, limited to the types of systems to which it can be applied. There have been some innovative methods for using molecular mechanics calculations in conjunction with quantum mechanics calculations for predicting electronic spectroscopic shifts by K. Coutinho et al. [40].

Electric Field Models

The second method used for incorporating solvent effects into theoretical calculations is based upon using an electric field to model the bulk solvent. The more popular methods use a dielectric continuum for simulating the solvent [41-58]. These methods work off of the assumption that the bulk properties obtained from the solution are the same as the specific interactions of the solvent molecules. These methods have been widely applied to quantum mechanics, since these methods for including solvent effects are essentially free when compared to the cost of calculating the solute molecule.

The dielectric continuum models have been shown to reproduce solvent effects for most systems that do not contain strong solute-solvent interactions [54] such as hydrogen bonds. This is a severe limitation since almost all solute-solvent interactions involving water have hydrogen bonding, and water is one of the most important solvents used in biological processes and several areas of chemistry. There have been attempts to parameterize the dielectric continuum models to properly account for the strong solute-solvent interactions [55-57]. Although these methods

reproduce experimental observations, they do not properly model the physics of the system. Another method that has been used for correcting dielectric continuum models for hydrogen bonding solvents is to include a small number of explicit water molecules in the quantum chemical calculation [58]. This technique works moderately well assuming a proper structure is obtained.

Hybrid Quantum Mechanics-Molecular Mechanics Methods

The idea of modeling the entire condensed media using quantum mechanics is not feasible and the use of the dielectric continuum model imposes a limitation upon the type of systems that can be studied. These restrictions of the current methods for calculating solvent effects has led to the development of hybrid quantum mechanics-molecular mechanics (QM/MM) methods. Warshel was the first to examine the division of the system into two regions: the first region is modeled using quantum mechanics, the second is modeled using molecular mechanics [59]. This method attempts to maintain all the advantages of quantum mechanics and molecular mechanics calculations while minimizing the disadvantages of each method [60,61].

The partitioning for the quantum mechanics and molecular mechanics regions is relatively simple for modeling solutions. The quantum mechanics region includes the solute molecule and may also include a few specific solvent molecules or the first solvation shell. The molecular mechanics region is the remainder of the solvent molecules not treated using quantum mechanics. The quantum mechanics solvent molecules are computationally expensive, and, therefore, a minimal number of them typically are included in the simulations.

This particular partitioning of the two regions maintains the ability of the quantum mechanics calculations to model charge transfer, covalent bond formation and dissociation and some solvent polarization. This method is also capable of modeling large numbers of solvent molecules, since they are treated using molecular mechanics.

The partitioning of the quantum mechanics system is not limited to modeling the whole solute molecule, although modeling part of the solute using quantum mechanics and the remainder using molecular mechanics introduces additional complications into the calculation. These additional complications arise from modeling half of a covalent bond using quantum mechanics while modeling the other half using molecular mechanics. The quantum mechanics region must be "capped" to properly model the electronic structure for the solute. There are several different models for "capping" the quantum mechanics regions, although the two most popular are modeling a hydrogen atom as the covalent bond or using a modified Fluorine atom to represent the "electrons" for the molecular mechanics region [62].

One of the specific goals of this work is to correctly model the environment of a solvated solute molecule. This is accomplished by using the hybrid quantum mechanics-molecular mechanics method to explicitly model the solvent molecules. The long range Coulombic interactions are included into this model through the use of the self-consistent reaction field (SCRF) method. The reaction field cavity size is chosen as the non-bonded cutoff range. The non-bonded cutoff is the distance that the explicit interactions are no longer calculated, since the forces and energies from the interactions separated by this distance or greater are assumed to be the same as those predicted by the bulk properties of the solution. The SCRF

model for modeling the long range interactions has the advantage over Ewald sums and fast multipole expansions that no artificial periodicity is built into the model and the bulk properties are correctly reproduced.

This new method for predicting the effects of solvation will then be applied to several test molecule sets to determine the reliability of this method for modeling hydrogen bonding complexes. The method is then applied to several chemical problems, such as investigating the origin of the large solvent shift observed in ruthenium (II) pentaaminopyridine spectra and investigating the origin of the solvent dependence of the phosphorescence and fluorescence properties of uracil.

CHAPTER 2 COMPUTATIONAL MODELS FOR LARGE SYSTEMS

Introduction

There are two types of large computational chemistry calculations. The first type of calculation models a small system, such as the reaction of CH_4 with H^+ , using a limited number of conservative approximations [63]. The second type of large calculation models a large system, such as modeling of the Trypsinogen-Trypsin transition using, several non-conservative approximations [64]; this is the kind of system I am interested in studying. When modeling the latter type of large system several complications arise from the inability to model the entire system that is observed experimentally. There are several competing methods for modeling large molecular systems using computational chemistry, some of which are described herein.

This chapter has a dual purpose; first, it provides a general review of some of the methods available for modeling large molecular systems, and second, it provides a common location for describing the exact methods I have implemented in the ZINDO program package [65]. This chapter is organized into three general areas: a description of the methods used for calculating potential energy, searching phase space, and modeling long range interactions and boundary effects.

Potential Energy Modeling

The type of potential energy calculation that is chosen significantly restricts the systems that can be studied. There are two methods commonly used for approximating the potential energy: molecular mechanics and quantum mechanics. Molecular mechanics calculations are able to model systems containing more than a hundred thousand atoms whereas quantum mechanics is limited to 30 atoms (*ab initio*) or a couple thousand atoms (semi-empirical). The major difference between the methods is that quantum mechanics explicitly models electrons and requires a solution to the Schrödinger equation. A quantum mechanics calculation is, therefore, capable of modeling systems that involve electronic phenomena such as covalent bond formation and dissociation and electronic spectroscopy. Another advantage of quantum mechanics is that it can be applied to any system without having to be parameterized.

Molecular Mechanics

Molecular mechanics, or force field, calculations represent an empirical model derived from general physics equations [66]. There are two different models which have been exploited in computational chemistry; all atom force fields and united atom force fields [67-75]. The all atom force fields treat every atom as a hard sphere, whereas the united atom force fields treat groups of atoms as hard spheres. The united atom force fields use chemical knowledge for determining what types of groups display similar chemical and physical properties, such as CH₂, CH₃, OH, and NH₂ groups. Regardless of the type of force field used, the hard spheres interact using the same model potentials.

The intramolecular interactions between the hard spheres is accomplished by connecting the specified spheres with springs to represent bonds (Figure 2.1).

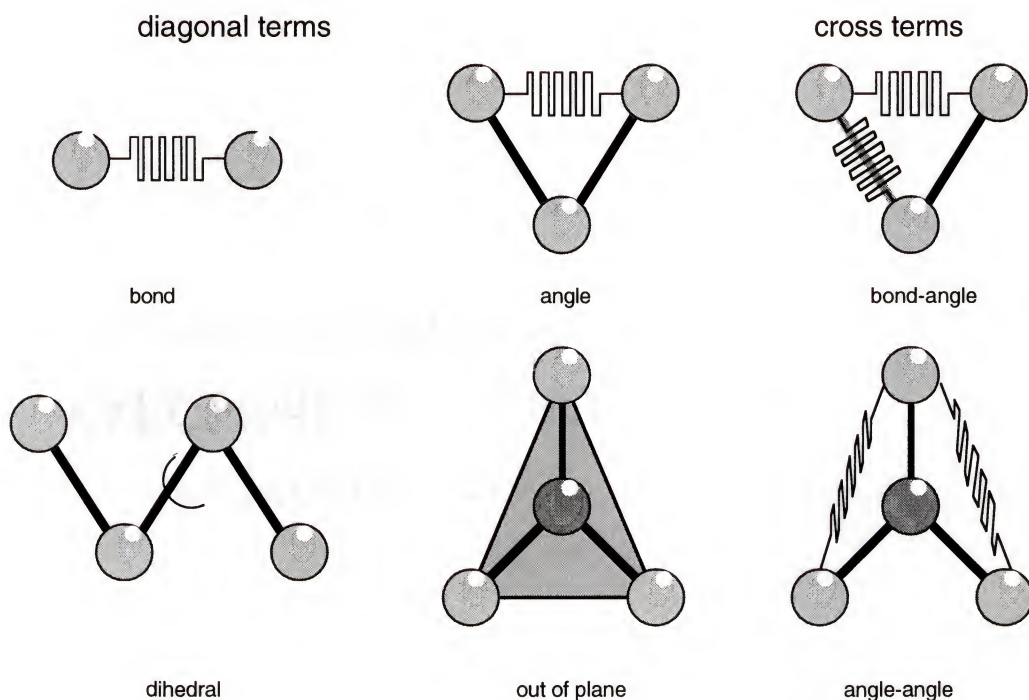


Figure 2.1: Schematic of Molecular Force Field Expressions.

Diagonal terms refer to interactions that can be expressed as a function of a single internal coordinate, whereas cross terms introduce coupled interactions involving two or more coordinates.

The Hamiltonian for the intramolecular interactions is

$$\begin{aligned} \hat{H}_{\text{intramolecular}} = & \sum_{\text{bonds}} \frac{K_r}{2} (r - r_o)^2 + \sum_{\text{angles}} \frac{K_\theta}{2} (\theta - \theta_o)^2 \\ & + \sum_{\text{dihedrals}} \sum_{n=1}^3 \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] . \end{aligned}$$

2.1

In equation 2.1, K_x are the spring force constants, x are the current value for the specified interactions (b , θ , ϕ), and x_o are the value for the measured property that minimizes the interactions (b_o , θ_o , ϕ_o).

The intermolecular molecular interactions are modeled using electrostatics. This may be accomplished using a Coulombic piece in addition to a Lennard-Jones (12-6) potential [76]. The Hamiltonian for the intermolecular interactions is

$$\hat{H}_{\text{intermolecular}} = \sum_{i \neq j} \frac{q_i q_j}{\epsilon R_{ij}} + \sum_{i \neq j} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^6 \right] \quad 2.2$$

In equation 2.2, q is the charge located at the center of each sphere, R is the distance between the centers of the two spheres, ϵ is the depth of the Lennard-Jones potential energy curve and σ is the Lennard-Jones parameter for determining the minimum energy distance between two spheres.

The molecular mechanics Hamiltonian is quite simple mathematically, and this method is also regarded as the “cheapest” level of theory which can be performed. The cost of this method scales as N_{MM}^2 , where N_{MM} is the number of hard spheres, with respect to the amount of memory required and the CPU time. Although this is the fastest method used in atomistic modeling computational chemistry, it is also the most restricted model since it does not have electrons and empirical data is

required to obtain the parameters. It can be made, however, to be quite accurate in reproducing geometry and thermochemistry.

Quantum Mechanics

From classical mechanics the energy of a system of interacting particles is the sum of the kinetic-energy (T) and potential-energy (V),

$$T + V = E \quad 2.3$$

Schrödinger suggested that the proper method for describing the wave character of particles was to replace the classical kinetic and potential energy functions with linear operators and then set up a wave equation [77,78],

$$\hat{H}\Psi = E\Psi \quad 2.4$$

where

$$\hat{H} = \hat{T} + \hat{V} \quad 2.5$$

The wavefunction (Ψ) describes the spatial motion of all the particles of the system moving in the field specified by the potential energy operator \hat{V} .

This initial observation of Schrödinger and Heisenberg [77-79] led the way for the first quantum mechanics calculation on a molecule in 1927 [80].

Although there are several techniques and approximate solutions for the Schrödinger equation, much of the work in this thesis is based upon the Self-Consistent Field (SCF) Molecular Orbitals (MO) methods [81-95]. The Hartree-Fock Self-Consistent Field (HF-SCF) method is used for geometries, whereas the configuration interaction (CI) method is used for spectroscopic results [96-100], for the remainder of this work.

Semi-Empirical Models

There are several different semi-empirical methods and parameter sets. The semi-empirical models are mostly based upon ignoring or approximating the two-electron integrals and using parameters to compensate for the neglected terms. The different types of two-electron integrals that are approximated by the semi-empirical methods can be seen by examining the closed-shell (RHF) the Hartree-Fock equation, the closed-shell limitation imposes no loss of generality.

$$F_{\mu\mu} = U_{\mu\mu} - \sum_{B \neq A} \langle \mu | \frac{Z_B}{R_B} | \mu \rangle + \sum_{\sigma, \lambda} P_{\sigma\lambda} \left(\langle \mu\sigma | \mu\lambda \rangle - \frac{1}{2} \langle \mu\mu | \sigma\lambda \rangle \right) \quad 2.6$$

where μ, σ, λ are the atomic orbitals (AO's), A and B are the atomic centers, and P is the atomic electron density. In Equation 2.6 each of the terms represents a different type of one- or two-electron integral. The first term is the core integral, which is an atomic-like integral,

$$U_{\mu\mu} = \left\langle \mu \left| \frac{-\nabla^2}{2} - \frac{Z_A}{R_A} \right| \mu \right\rangle + V \quad 2.7$$

The second term of Equation 2.6 involves one-electron two-center nuclear attraction integrals $\langle \mu | \frac{Z_B}{R_B} | \mu \rangle$, (Figure 2.2).

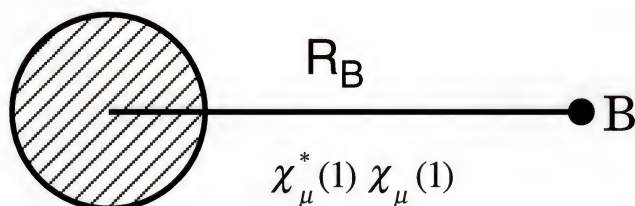


Figure 2.2: One-Electron Two-Center Integral

The next term in Equation 2.6 involves the two-center two-electron integrals,

$$\langle \mu\sigma | \nu\lambda \rangle = \int d\tau(1) \int d\tau(2) \chi_\mu^*(1) \chi_\nu(1) \frac{1}{r_{12}} \chi_\sigma^*(2) \chi_\lambda(2) \quad 2.8$$

Examining these terms further, $\langle \mu\nu | \mu\nu \rangle$ is shown in Figure 2.3.

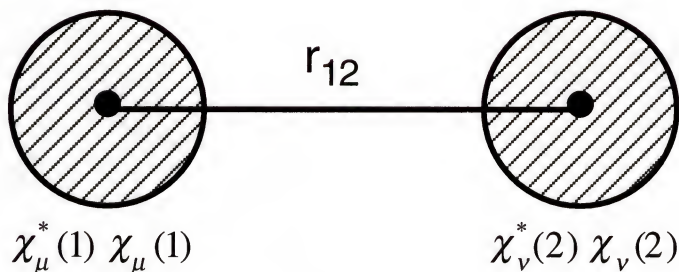


Figure 2.3: One-Electron Two-Center Integral

Examining more of the two-electron terms of Equation 2.6, we sketch the differential atomic and diatomic overlap integrals in Figure 2.4.

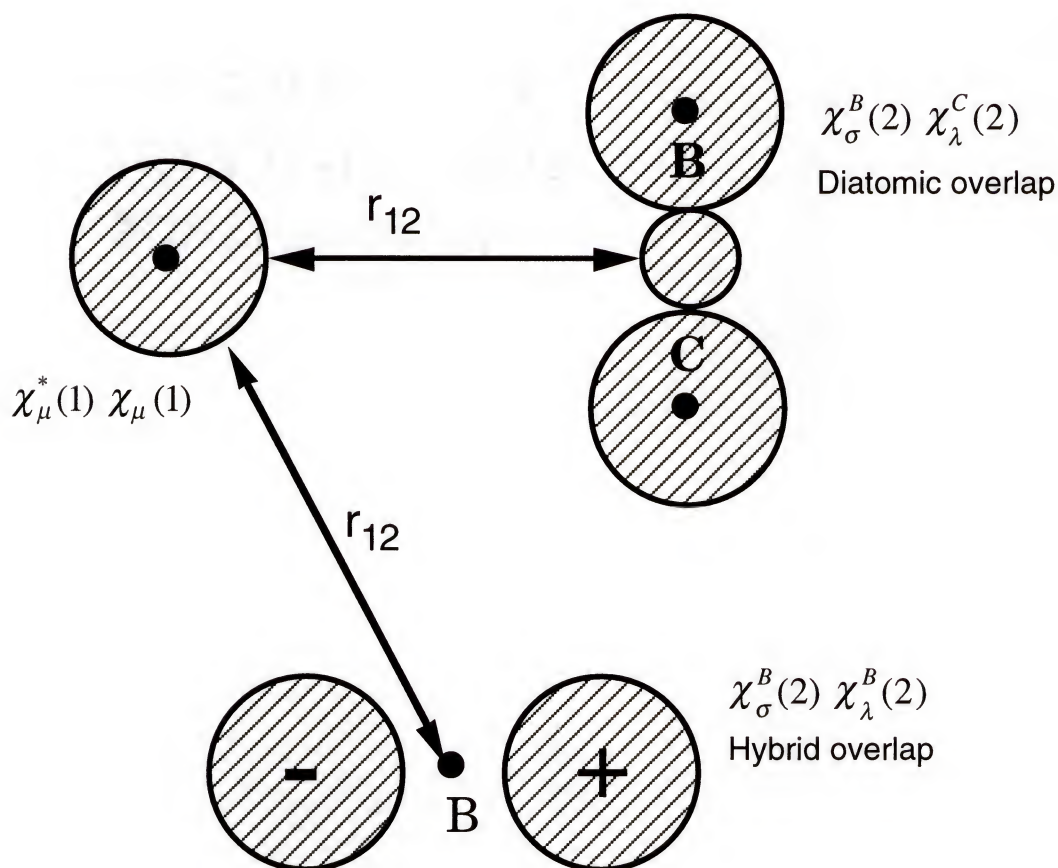


Figure 2.4: Two-Electron Two-Center Integral Differential Atomic and Diatomic Overlap

The differential diatomic overlap integrals are ignored by all the semi-empirical methods, while the Neglect of Differential Diatomic Overlap (NDDO) model includes the differential atomic overlap integrals. The computational limitation for quantum mechanics is the matrix

diagonalization for the solution of the secular equation, which is an N_{QM}^3 problem, where N_{QM} is the number of basis functions. All MO methods are limited by the N_{QM}^3 matrix problem, although there are some methods that reduce the N_{QM}^3 problem to several mM_{QM}^3 problems, where m is the number of subset problems to be solved, and M_{QM} is the number of basis functions in that partition. These techniques have been to scale as $N_{QM} \log[N_{QM}]$ overall [101-103].

The zero differential overlap (ZDO) [104-108] methods systematically exclude all the three- and four-center integrals, which reduces the number of integrals to be calculated to N_{QM}^2 . The remaining integrals that remain are calculated, or parameterized from model calculations, or directly from experiment. In addition to the reduction of the integrals, the ZDO approximation also requires utilization of a minimum basis set representation for atoms. This representation also reduces the number of integrals that have to be calculated. There is also a core-orbital valence-orbital separation; representing the core through parameters also reducing the number of integrals since the core orbitals are replaced with a function.

There are three different common levels of ZDO approximation; Complete Neglect of Differential Overlap (CNDO), Intermediate Neglect of Differential Overlap (INDO), and Neglect of Differential Diatomic Overlap (NDDO). The CNDO model has the most severe approximations. The CNDO [109] scheme replaces

$$\chi_{\mu}^A(1)\chi_{\nu}^A(1)d\tau(1) \rightarrow \delta_{\mu\nu}\chi_{\mu}^A(1)\chi_{\mu}^A(1)d\tau(1) \quad 2.9$$

whenever it occurs in an integral over a spherically symmetric operator. The superscripts A and B refer to the atomic centers, and μ and ν refer to individual orbitals. The bars over the subscript indicate that the actual orbital χ_μ is replaced by an s symmetry orbital of the same spatial extent. This condition is required to maintain the rotational invariance. Using this approximation the metrix, or overlap matrix, Δ , becomes the unit matrix,

$$\Delta = \mathbf{1} \quad 2.10$$

and

$$F_{\mu\mu}^{AA} = U_{\mu\mu} + \sum_{\nu} P_{\nu\nu} \langle \overline{\mu\nu} | \overline{\mu\nu} \rangle - P_{\mu\mu} \langle \overline{\mu\mu} | \overline{\mu\mu} \rangle / 2 - \sum_{B \neq A} Z_B \langle \overline{\mu} | 1 / R_B | \overline{\mu} \rangle \quad 2.11$$

$$F_{\mu\nu}^{AB} = H_{\mu\nu} - P_{\mu\nu} \langle \overline{\mu\nu} | \overline{\mu\nu} \rangle / 2 \quad 2.12$$

The integral $U_{\mu\mu}$ is called the core integral and is defined as the atomic-like one electron integral, Equation 2.7. The role of V , in Equation 2.7, is to keep the valence electrons from penetrating into the inner-shell regions. If U is obtained from an empirical procedure, or from atomic information on valence-electrons only methods, then the one-center part of V is included in this parameterization in a natural fashion [110-112].

The INDO model [113] contains all of the terms that were included in the CNDO model, and all of the one-center two-electron integrals as well. The INDO Fock matrix for a closed shell is

$$F_{\mu\mu}^{AA} = U_{\mu\mu} + \sum_{s,t}^A P_{st} (\langle \mu s | \mu t \rangle - \langle \mu s | t \mu \rangle / 2) +$$

$$\sum_{B \neq A}^B \sum_s P_{ss} \langle \overline{\mu s} | \overline{\mu s} \rangle - \sum_{B \neq A} Z_B \langle \overline{\mu} | 1/R_B | \overline{\mu} \rangle \quad 2.13$$

$$F_{\mu\nu}^{AA} = H_{\mu\nu} + \sum_{s,t}^A P_{st} (\langle \mu s | \nu t \rangle - \langle \mu s | t \nu \rangle / 2) \quad 2.14$$

$$F_{\mu\nu}^{AB} = H_{\mu\nu} - P_{\mu\nu} \langle \overline{\mu\nu} | \overline{\mu\nu} \rangle / 2 \quad 2.15$$

The formation of the Fock matrix scales as an N_{QM}^2 step, the same as does CNDO. The INDO model is better than the CNDO model, as would be expected from the inclusion of additional integrals. This improvement manifests itself in the prediction of better angles, although there is not a significant improvement. The INDO model does predict molecular electronic optical properties much better than the CNDO model [114], especially for transition metal complexes where the additional one-center two-electron terms can be large.

The neglect of differential diatomic overlap (NDDO) model [115] is derived from the following replacement:

$$\chi_{\mu}^A(1) \chi_{\nu}^B(1) d\tau(1) \rightarrow \delta_{AB} \chi_{\mu}^A(1) \chi_{\nu}^A(1) d\tau(1) \quad 2.16$$

The Fock matrix is then given by

$$F_{\mu\mu}^{AA} = U_{\mu\mu} + \sum_{s,t}^B P_{st} [\langle \mu s | \mu t \rangle - \langle \mu s | t \mu \rangle / 2] - \sum_{B \neq A} Z_B \langle \mu | 1/R_B | \mu \rangle \quad 2.17$$

$$F_{\mu\nu}^{AA} = \sum_B \sum_{s,t}^B P_{st} \langle \mu s | \nu t \rangle - \sum_{s,t}^A P_{st} \langle \mu s | t \nu \rangle / 2 - \sum_B Z_B \langle \mu | 1/R_B | \nu \rangle \quad 2.18$$

$$F_{\mu\nu}^{AB} = H_{\mu\nu} - \sum_s^A \sum_t^B P_{st} \langle \mu t | s \nu \rangle / 2 \quad 2.19$$

There are some important advantages that the NDDO model has over the CNDO or INDO models. The NDDO model includes orbital anisotropies, and therefore the requirement for spherically averaging two-center two-electron integrals is removed. The NDDO terms are also generated in a natural manner, and it appears that this model is easier to parameterize than either the CNDO or INDO models [116-118].

Hybrid Quantum Mechanics-Molecular Mechanics Models

The idea for combining quantum mechanics and molecular mechanics (QM/MM) was first proposed by Warshel [59]. Since then there have been several other hybrid QM/MM models proposed [119-124]. These

hybrid QM/MM models have been applied to various chemical and biochemical problems including solvent effects and protein modeling. The hybrid QM/MM model is best understood by using an “onion” model representation, (Figure 2.5).

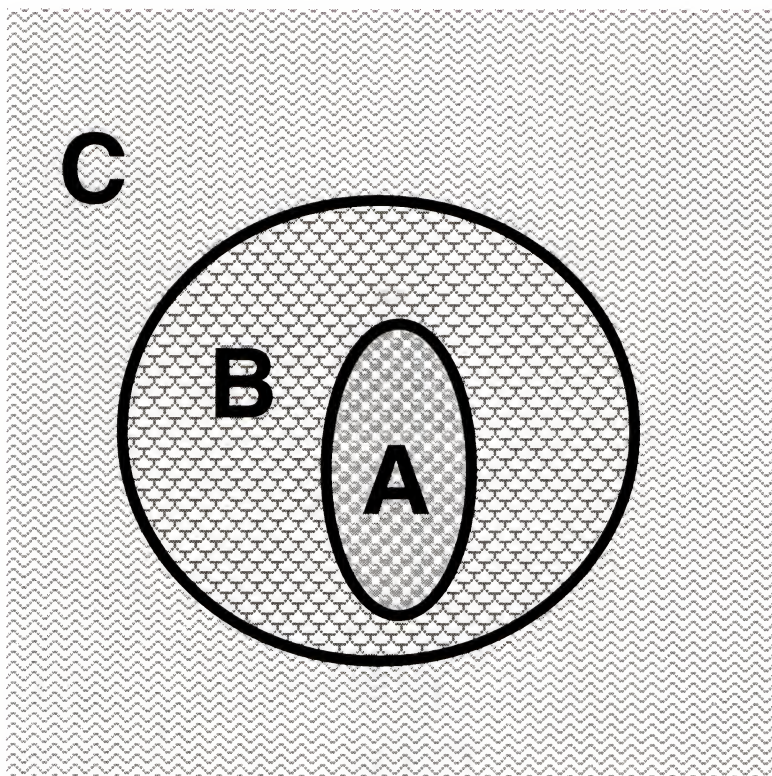


Figure 2.5: Pictorial View of QM/MM Model

Where region “A” is treated using any quantum mechanics model, region “B” is treated using classic mechanics model, and region “C” is treated using a bulk dielectric, Ewald summation, or fast multipole approximation.

The mathematical derivation of the specific QM/MM model that I have implemented starts with the creation of a Hamiltonian that can then

be solved by using the methods outlined in the quantum mechanics section. The new Hamiltonian is constructed from three independent partitions, and a Hamiltonian that is the sum of three parts.

$$\hat{H}_{Total} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} \quad 2.20$$

The first partition is treated using the standard quantum mechanics Born-Oppenheimer fixed nuclei Hamiltonian [81],

$$\hat{H}_{QM} = -\frac{1}{2} \sum_i^N \nabla_i^2 + \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}} - \sum_i^N \sum_A^m \frac{Z_A}{r_{iA}} + \sum_A^m \sum_{B>A}^m \frac{Z_A Z_B}{R_{AB}} \quad 2.21$$

where i and j are electrons, A and B are nuclei, Z is the atomic number, and N and m are the number of electrons and nuclei, respectively. The second partition is defined by the molecular mechanics Hamiltonian, (see MM section for details)

$$\begin{aligned} \hat{H}_{MM} = & \sum_{bonds} \frac{K_r}{2} (r - r_o)^2 + \sum_{angles} \frac{K_\theta}{2} (\theta - \theta_o)^2 \\ & + \sum_{dihedrals} \sum_{n=1}^3 \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i \neq j} \frac{q_i q_j}{\epsilon R_{ij}} \\ & + \sum_{i \neq j} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^6 \right] \end{aligned} \quad 2.22$$

While the final partition is designed to allow the molecular mechanics and quantum mechanics to interact

$$\begin{aligned}\hat{H}_{QM/MM} = & -\sum_i \sum_s \frac{q_s}{r_{is}} - \sum_A \sum_s \frac{Z_A q_s}{R_{As}} \\ & + \sum_A \sum_s 4\epsilon_{As} \left\{ \left(\frac{\sigma}{R_{As}} \right)^{12} - \left(\frac{\sigma}{R_{As}} \right)^6 \right\}\end{aligned}\tag{2.23}$$

where the classical mechanics atoms are represented by s , the electrons by i , and the quantum mechanics nuclei by A .

It is now convenient to invoke the Born-Oppenheimer approximation [81] on the Hamiltonian (\hat{H}_{Total}), which simplifies the Hamiltonian yielding an electronic energy that provides a potential energy for the nuclei,

$$\hat{H}_{Total} = \hat{H}_{electronic} + E_{nuclear}\tag{2.24}$$

where $H_{electronic}$ is

$$\hat{H}_{electronic} = \frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_A \frac{Z_A}{R_{Ai}} + \sum_{i<j} \frac{1}{r_{ij}} - \sum_i \sum_s \frac{q_s}{r_{is}}\tag{2.25}$$

and $E_{nuclear}$ is

$$\begin{aligned}
E_{nuclear} = & \sum_{A < B} \frac{Z_A Z_B}{R_{AB}} + E_{MM} - \sum_A \sum_s \frac{Z_A q_s}{R_{As}} \\
& + \sum_A \sum_s 4\epsilon_{As} \left\{ \left(\frac{\sigma}{R_{As}} \right)^{12} - \left(\frac{\sigma}{R_{As}} \right)^6 \right\}
\end{aligned} \tag{2.26}$$

Where E_{MM} is the potential energy of the molecular mechanics region. The energy for the Hamiltonian (\hat{H}_{Total}) then becomes,

$$E_{Total} = \frac{\langle \Phi_{electronic} | H_{electronic} | \Phi_{electronic} \rangle}{\langle \Phi_{electronic} | \Phi_{electronic} \rangle} + E_{nuclear} \tag{2.27}$$

Where E_{Total} is the potential energy for the whole system which is dependent upon the nuclear coordinates.

One of the more interesting subtleties of this formalism for the Hamiltonian is that the electronic wave function is only dependent upon one piece of $H_{QM/MM}$. This dependence was not initially accounted for in any of the QM/MM models developed until we were testing for molecular indistinguishability for the water dimer using our Hybrid QM/MM model (see chapter 3). In addition, for dynamics and geometry optimization techniques that require the calculation of forces, this piece must be totally imbedded (no partitioning as others have done) in the QM region to avoid having to use a coupled-perturbed Hartree-Fock theory.

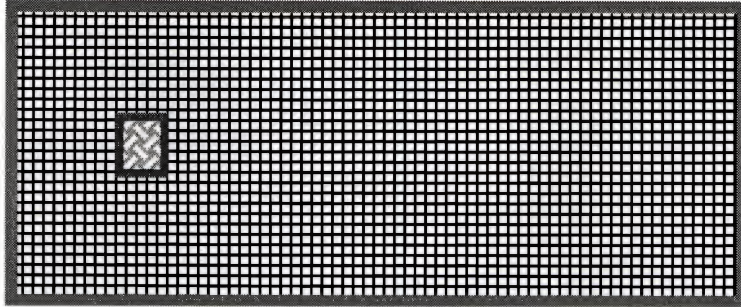
This QM/MM Hamiltonian is also unique in that a self-consistent reaction field is used to model the long range coulombic interactions. The dielectric continuum is placed on the molecular cutoff limit, usually 12-15

Å. The advantage of using a reaction field versus a particle mesh Ewald or fast multipole method is that the dielectric continuum exactly models the bulk solvent, whereas the other two methods are based upon a sample cell that may not be representative of the bulk solvent. In addition the periodic nature of the latter two methods also introduces some periodicity into the bulk solvent.

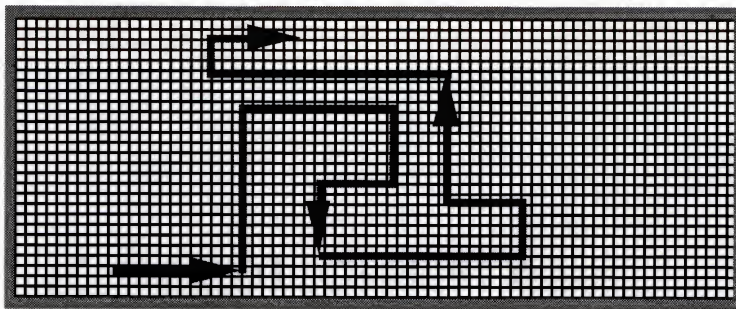
Geometric Modeling Strategies

There are several different techniques that can be employed for searching phase space. These techniques are generally divided into two classes: stochastic methods and minimization techniques. There are several differences between the two classes. The stochastic methods can be used to calculate thermodynamic properties as they search phase space, whereas the minimization techniques are efficient methods for finding local minima. Figure 2.6 is a pictorial view of how these various techniques search phase space. This figure also shows the relative amount of phase space that these methods are searching. In Figure 2.6 the large boxes represent phase space. The geometry optimization figure shows that these techniques do not search phase space, but find the nearest local energy minimum. The space searched using dynamics is represented by the line winding through phase space. This line is continuous and the path is determined from solving Newton's equations of motion. The area of phase space covered by Monte Carlo techniques is indicated by the checkered pattern regions. The area searched is not continuous, since Monte Carlo techniques randomly search phase space in a time independent manner.

Geometry Optimization



Dynamics



Monte Carlo

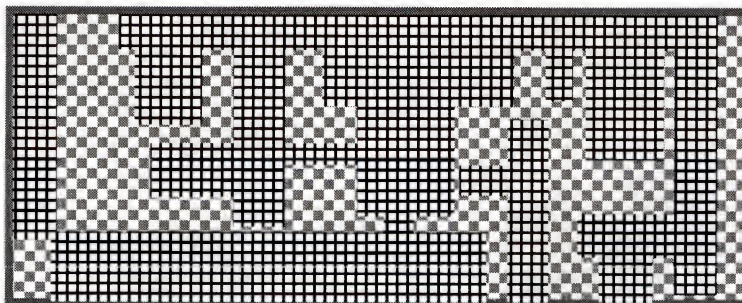


Figure 2.6: Graphical Representation of how the Various Methods Search Phase Space.

The Boxes represent phase space while the lines and textured boxes represent how the various techniques search phase space.

The dynamics and Monte Carlo techniques will both eventually span all of the phase space if an infinitely long simulation is performed, although this is not feasible except for the most simple cases.

Geometry Optimization

Geometry optimization techniques are of vital importance for both molecular mechanics and quantum mechanics calculations. These techniques are capable of finding local minima on the potential energy surfaces and provide starting structures for transition state searches and stochastic methods. There are several different techniques available for finding these local minima. These geometry optimization techniques are limited in that they will usually find the minima nearest to the starting geometry as opposed to a global minima. In spite of this drawback these methods are still the most powerful methods available for finding local minima on the potential energy surface [125].

Steepest Descent

The steepest descent method [126,127] is classified as a first order minimizer since it only uses the first derivatives of the potential energy for predicting the next step. This method moves along the steepest slope of the interatomic forces on the potential energy surface. This method scales the step that each atom makes by the negative of the derivative of the potential energy, as shown in Figure 2.7. The steepest descent method is an effective method for alleviating large forces on atoms. This technique is usually used to obtain a quasi-optimized structure from which a better optimization technique can be employed. The major disadvantage of this method is that the speed of convergence is highly dependent on the step size.

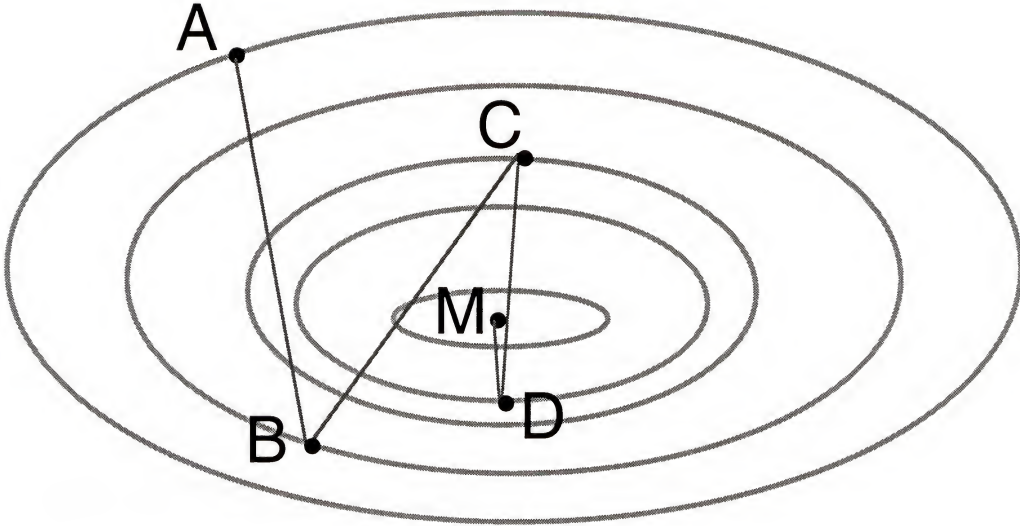


Figure 2.7: Steepest Descent Method

M is the minimum point on the potential energy surface, point A is the starting point, points B, C and D are intermediate steps. This is an example of a successful steepest descent search, if the step size chosen is too large this technique converges very slowly if at all.

Conjugate Gradient

The conjugate gradient method [128] differs from the steepest descent in that it uses both the current gradient and the previous search direction to predict the new step. The conjugate gradient method is also considered a first order method since it only uses the gradient. The conjugate gradient method has the advantage that it uses the minimization history to calculate the new search direction. This method therefore converges faster than the steepest descent method (Figure 2.8). The conjugate gradient method also provides a scaling factor for determining the step size. The step sizes used in the conjugate gradient method are more optimized than those obtained using a steepest descent method.

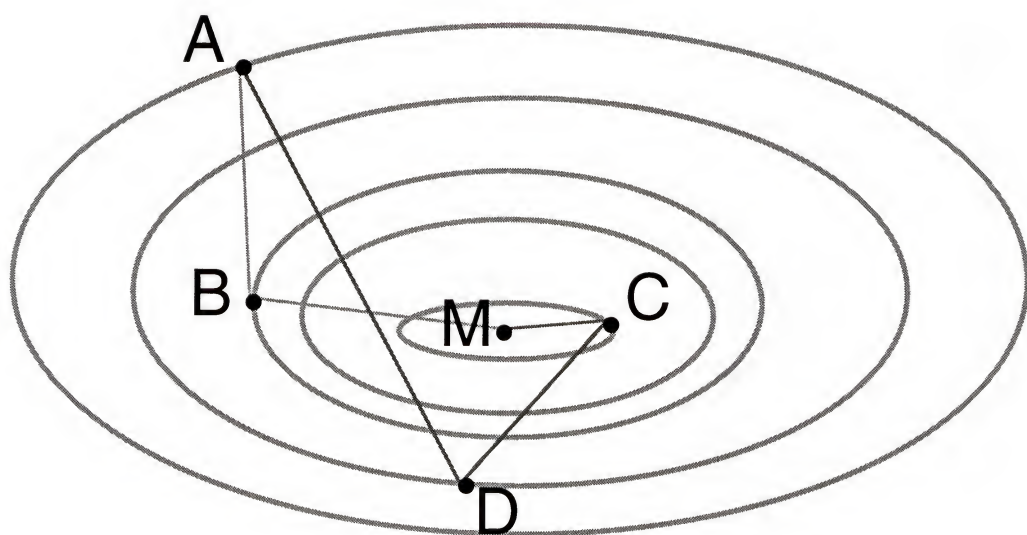


Figure 2.8: Conjugate Gradient

Comparing the conjugate gradient with the steepest descent method. This system can reach the minimum using only two steps (A-B-M). If the first step is too large this method still converges (A-D-C-M), whereas the steepest descent method may not converge.

Quasi-Newton Update Methods

The quasi-Newton methods are a second order method, since they require analytical gradients and then build up the second derivative, or the Hessian, as they move along the hypersurface [129]. There are several different methods that are frequently used to update the Hessian, such as Murtagh-Sargent (MS), Broyden-Fletcher-Goldfarb-Shanno (BFGS), and Davidson-Fletcher-Powell (DFP) methods [130-135]. For a more detailed description of the quasi-Newton techniques and other geometry optimization techniques see reference [136-138].

Molecular Dynamics

Molecular dynamics is used for searching phase space in a systematic random manner. This can be seen by a simple analogy to methods that can be used for dealing cards. Think of a deck of cards, if you are wanting to give four players in a poker game a random collection of five cards, you can just randomly pick five cards for each player (Monte Carlo) or you could choose a starting card and follow a systematic procedure for picking the next card (Molecular Dynamics). Both of the techniques described will give the four players a random assortment of cards from the deck, which is analogous to the difference in how the Monte Carlo and molecular dynamics methods search phase space.

Molecular dynamics uses Newton's second law of motion (2.26) for determining how the next set of phase space coordinates is chosen [139,140].

$$\vec{F} = m\vec{a} \tag{2.28}$$

Where \vec{F} is the force, m is the mass, and \vec{a} is the acceleration or the second derivative of the position with respect to time. While an analytical solution is available for Newton's second law for modeling the interactions of hard spheres, using a more realistic model for the potential energy which varies continuously with respect to the distance of separation, requires a numerical method for solving the differential equations. This is usually achieved by using finite-difference methods to replace the differentials dx and dt with finite differences Δx and Δt such that the position after a

specified time Δt is given by $x(t_o + \Delta t)$. The finite-difference methods that are described herein originate from a truncated Taylor series expansion.

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt} \Delta t + \frac{1}{2!} \frac{d^2 x(t)}{dt^2} (\Delta t)^2 + \frac{1}{3!} \frac{d^3 x(t)}{dt^3} (\Delta t)^3 + \dots \quad 2.29$$

For simplicity, this series will be truncated after the first term for the remainder of the molecular dynamics development; this is also known as Euler's method.

$$x(t + \Delta t) = x(t) + \dot{x}(t) \Delta t \quad 2.30$$

This provides a method for estimating the value of x after a specified time by extrapolating a straight line that has a slope dx/dt evaluated at t . The finite-difference method introduces two types of error: truncation error and round-off error. The truncation error (te) is defined as the error associated with how the truncated Taylor series approximates the true system. This is also thought of as the precision of the calculation. This error is measured by the first term of the Taylor series that has been omitted, such that the truncation error associated with the Euler's method is

$$te = \frac{1}{2} \frac{d^2 x(t)}{dt^2} (\Delta t)^2 \quad 2.31$$

Euler's method is first order, since an n -th order equation is said to have a truncation error that varies as $(\Delta t)^{n+1}$. The truncation error is a constant for a given method and the size of this error is reduced by using higher order equations. The second kind of error is the round-off error which encompasses all of the errors that results from the implementation of the finite-difference algorithm.

There are two methods commonly used by finite-difference molecular dynamics simulations, the first is a third-order Störmer algorithm using the Verlet method of combining two Taylor series expansions [141,142]. Consider the Taylor series for position from time (t) forward to $t + \Delta t$:

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt}(\Delta t) + \frac{1}{2!} \frac{d^2 x(t)}{dt^2}(\Delta t)^2 + \frac{1}{3!} \frac{d^3 x(t)}{dt^3}(\Delta t)^3 + O(\Delta t)^4 \quad 2.32$$

Then write the Taylor series from t backward to $t - \Delta t$:

$$x(t - \Delta t) = x(t) - \frac{dx(t)}{dt}(\Delta t) + \frac{1}{2!} \frac{d^2 x(t)}{dt^2}(\Delta t)^2 - \frac{1}{3!} \frac{d^3 x(t)}{dt^3}(\Delta t)^3 + O(\Delta t)^4 \quad 2.33$$

Adding equation 2.32 and 2.33 eliminates all odd-order terms, yielding

$$x(t + \Delta t) + x(t - \Delta t) = 2x(t) + \frac{1}{2!} \frac{d^2 x(t)}{dt^2} \Delta t^2 + O(\Delta t^4) \quad 2.34$$

This is Verlet's algorithm that relates the new position to the two previous positions. The truncation error varies as $((\Delta t)^4)$ and is, therefore, third order, even though it does not contain any third order derivatives. This algorithm also provides a method for determining the position without having to calculate any velocities. The acceleration is calculated from Newton's second law (Equation 2.28). The velocities in practice can be calculated using several methods; the most common of these are estimating the velocities at half-steps:

$$v(t + \frac{1}{2} \Delta t) \approx \frac{x(t + \Delta t) - x(t)}{\Delta t} \quad 2.35$$

or by using the first-order central difference estimator

$$v(t) \approx \frac{x(t + \Delta t) - x(t - \Delta t)}{2 \Delta t} \quad 2.36$$

The Verlet algorithm is a two-step method and is not self-starting since it requires more than the initial positions (x) and velocities (v) to begin a simulation. This is usually overcome by using a simple Euler model to calculate the information at $t - \Delta t$ to start a simulation. This algorithm offers good stability using a large time step, although it does favor the positions more than the velocities, which could bias the phase space which this method spans.

Another popular method for solving the finite-difference equations in molecular dynamics is a predictor-corrector algorithm [143]. These

methods involve a three-step process for each set of phase space coordinates:

1. Predict the position $x(t + \Delta t)$ and velocity $v(t + \Delta t)$ at the end of each step.
2. Evaluate the forces at $t + \Delta t$ using the predicted position.
3. Correct the predictions using a combination of the predicted and calculated values of position and velocity.

The third-order Adams-Bashforth predictor [143] is given by

$$\vec{r}_{n+1} = \vec{r}_n + \frac{(23\vec{v}_n - 16\vec{v}_{n-1} + 5\vec{v}_{n-2})}{12}(t_{n+1} - t_n) \quad 2.37$$

coupled with the Adams-Moulton correctors [143]:

$$\vec{v}_{n+1} = \vec{v}_n + \frac{(5\vec{a}_{n+1} + 8\vec{a}_n - \vec{a}_{n-1})}{12}(t_{n+1} - t_n) \quad 2.38$$

$$\vec{r}_{n+1} = \vec{r}_n + \frac{(5\vec{v}_{n+1} + 8\vec{v}_n - \vec{v}_{n-1})}{12}(t_{n+1} - t_n) \quad 2.39$$

This combination seemed to provide good stability for a relatively cheap computational cost. Therefore I have chosen to use this method for the molecular dynamics simulations.

Another important aspect of molecular dynamics is a method for regulating the temperature. This is crucial for both the canonical

ensemble (NVT) and the isothermal-isobaric ensemble (NPT). There are several different procedures for regulating the temperature in molecular dynamics simulations [144-148]. I have chosen to implement a method that couples the system to an external heat bath [144]. This is accomplished by introducing a friction term into the equations of motion,

$$m_k \vec{a}_i = \vec{f}_k - m_k \gamma \vec{v}_k + R_k(t) \quad 2.40$$

where f_k is the systematic forces and R_k is the Gaussian stochastic variable that is required to have a zero mean value. The intensity of the stochastic Gaussian variable is given by

$$\langle R_k(t) R_l(t + \tau) \rangle = 2m_k \gamma_k k_B T_o \delta(\tau) \delta_{kl} \quad 2.41$$

where the damping constant (γ_k) determines the magnitude of the coupling between the heat bath and the system. The time dependence for the temperature scaling is then obtained from the derivative of the total kinetic energy (E_{kin}):

$$\frac{dE_{\text{kin}}}{dt} = \lim_{\Delta t \rightarrow 0} \left[\sum_{i=1}^{3N} \frac{m_i \vec{v}_i^2(t + \Delta t) - m_i \vec{v}_i^2(t)}{2} / \Delta t \right] \quad 2.42$$

where N is the number of particles. From these equations the time dependent temperature scaling factor (λ) can be derived, such that,

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left(\frac{T_o}{T} - 1 \right) \right]^{1/2} \quad 2.43$$

where λ is used to scale the velocities, $\vec{v} = \lambda \vec{v}$ and τ_t is the heat bath friction coefficient. The derivation of λ from equations (2.40 - 2.42) is given in the Berendsen paper [144]. This particular temperature scaling model minimizes the difference in the kinetic energy derived from the average square of the velocities,

$$\sum_i \frac{m_i (\Delta v_i)^2}{2} \quad 2.44$$

while the change in kinetic energy,

$$\sum_i \Delta \left(\frac{m_i v_i^2}{2} \right) \quad 2.45$$

is constrained to a specified value. A similar type of scaling may be performed on the pressure, if the ensemble allows pressure fluctuations, such as the NPT ensemble. The pressure scaling is then given by,

$$\mu = \left[1 - \beta_T \frac{\Delta t}{\tau_P} (P - P_o) \right] \quad 2.46$$

where μ is used to scale the volume of the system, $V = \mu V$ and τ_p is the pressure friction coefficient. This equation relates the pressure change to the isothermal compressibility (β_T).

Brownian Dynamics

Molecular dynamics is the predominate dynamics method used for studying chemical systems. The Brownian dynamics model is a viable method to use when the interest is limited to a small portion of a large system and the molecular details of the remainder of the system are not of concern. This type of problem is commonly encountered when studying solvent effects, since it is the first two to three solvation shells that directly interact with the solute. This type of dynamics has the advantage that smaller systems can be modeled, which allows longer simulation times. In Brownian dynamics, the forces acting on the solute molecule have a component which arises from the intermolecular interactions in the solute and/or any external field, a component arising from the solvent friction, and a third random component to model the thermal fluctuations of the solvent molecules:

$$m \left(\frac{d\vec{v}}{dt} \right) = \vec{F} - \zeta \vec{v} + R(t) \quad 2.47$$

where \vec{F} represents the forces arising from the explicit intermolecular and intramolecular interactions of the solute and ζ is the friction coefficient. As

before, the random forces $R(t)$ are chosen in such a manner that they have a zero mean value and a rapidly decaying autocorrelation function [149-152].

Monte Carlo

The Monte Carlo sampling method was developed by Ulam and Metropolis at the end of the Second World War to study diffusion of neutrons in fissionable media [153]. Metropolis and Ulam were not the first people to use sampling experiments to investigate such problems, earlier uses of these techniques were examined by statisticians at the turn of the century [154]. The innovation between what Ulam and Metropolis did was to realize that determinate mathematical problems could be treated by finding a probabilistic analogue which could then be solved using stochastic sampling. Metropolis in 1953 was the first to use the Monte Carlo method in a computational chemistry calculation [155].

The probability for a system being in a given microstate is given by the Boltzmann weighting factor

$$\rho(\alpha) = \frac{\exp[-H(\alpha) / kT]}{Z} \quad 2.48$$

where Z is a partition function given by

$$Z = \sum \exp\left[\frac{-H(\alpha_i)}{kT}\right] \quad 2.49$$

and H represents the Hamiltonian without the kinetic energy terms. The phase space configuration $(\rho(\alpha_i))$ must satisfy the principle of detailed balance. That is to say if p_{ij} represents the probability for a transition from state i to state j , then

$$\rho(\alpha_i)p_{ij} = \sum_i \rho(\alpha_j)p_{ji} \quad 2.50$$

such that ρ also satisfies the steady state condition

$$\rho(\alpha_j) = \sum_i \rho(\alpha_i)p_{ji} \quad 2.51$$

This allows any average thermodynamic property $\langle A \rangle$ of the system to be given by

$$\langle A \rangle = \frac{\sum A_i \exp[-H(\alpha_i) / kt]}{Z} \quad 2.52$$

Since all of the sums are over an infinite number of phase space points, the evaluation of Equation 2.50 is not computationally feasible. Metropolis et al. [155] formulated that these infinite sums can be replaced by an average over a finite number of points, which are selected according to their Boltzmann weighting factors. This is achieved by generating a new configuration from the current configuration with a transition probability

equal to the ratio of their Boltzmann weighting factors. The use of this procedure removes the requirement of *a-priori* knowledge of the partition function.

In practice a simulation is done in the following manner [156]. An initial set of atomic coordinates for the system is chosen as the initial state. These atoms interact through the specified potential energy function. A random Monte Carlo move is attempted to generate a new configuration. This movement can be an atomic displacement, molecular displacement, molecular rotation, or any other move applicable to the system. The probability for the transition from configuration i to the new configuration j is then calculated as

$$p_{ij} = \exp\left[\frac{-\Delta H}{kT}\right] \quad 2.53$$

where

$$\Delta H = H_{new} - H_{old} \quad 2.54$$

If $\Delta H < 0$, then $P_{ij} = 1$, the move is accepted. This means that the new configuration is lower in energy than the old configuration. The new configuration is now used as the initial state and another random Monte Carlo move is made. This process is repeated until the specified number of configurations has been accepted. If $\Delta H > 0$, then $p_{ij} < 1$, and the specified move must be tested. A random number (z) between 0 and 1 is generated

and then compared with p_{ij} . If $p_{ij} > z$, the move is accepted and the current configuration is taken as the new configuration; otherwise the move is rejected. The structural and other static properties are averaged over the accepted configurations. The accepted configurations are the result of a random walk through phase space which forms a Markov chain [157,158]. Although the Monte Carlo simulations are time independent, dynamical properties can be observed by associating each new configuration to a finite time. Although this artificial time will produce proper time averaged properties, the steps clearly do not represent the true time evolution of the physical system.

The Metropolis Monte Carlo method does not specify what types of moves that can be made. This is another point that most developers of these methods neglect to specify in their publications. This lack of available information led us to experiment with various types of moves. The objective of a Monte Carlo simulation when modeling chemical systems is to span phase space. This is achieved by making random moves until a defined number of uncorrelated structures is obtained. These uncorrelated structures are then used for determining the averaged bulk physical properties. Traditionally the Monte Carlo method has been applied to classical mechanics systems where the potential energy calculation scales as the number of atoms squared. This relatively fast computation of the potential energy enables inefficient Monte Carlo step schemes to exist, since calculating the potential energy $>1,000$ per uncorrelated structure is computationally feasible for most systems using classical mechanics. The implementation of quantum mechanics with its N_{QM}^3 or higher dependence on the number of basis functions makes this type of Monte Carlo simulation unfeasible.

A new form of Monte Carlo simulations is, therefore, needed when simulating systems that contain quantum mechanics. This new scheme is an optimization of the Monte Carlo method. Instead of calculating the potential energy 1,000 times for one uncorrelated structure, we are trying to get an uncorrelated structure for every 50 potential energy calculations. There are two types of Monte Carlo steps used in the classical calculations.

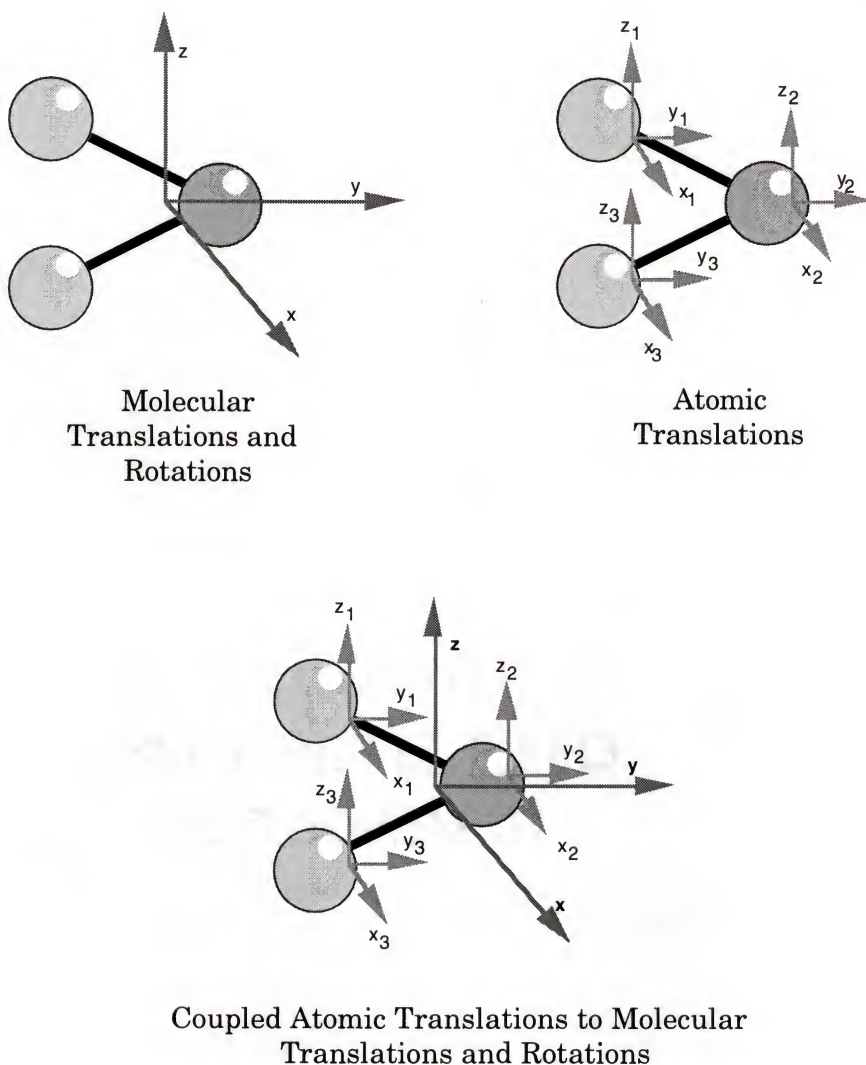


Figure 2.9: Types of Monte Carlo Steps

The first type involves rigid molecules; this type of simulation is based upon molecular translations, and sometime, rotations. The second type of simulation involves atomic translations. The first type of simulation is usually used for large systems involving many rigid molecules, such as solvent molecules, whereas the second type is usually used for studying the properties large molecule systems, such as vibrational structure. The new Monte Carlo method for simulating systems using quantum mechanics combines these two types of simulations. The new method we use employs both molecular and atomic translations coupled with molecular rotations. The various types of Monte Carlo steps described are shown in Figure 2.9.

The new type of Monte Carlo steps was tested against the rigid solvent molecule model, molecular rotations and translations only. The test was performed upon a box of 98 molecular mechanics SPC water molecules. Each of the Monte Carlo methods generated 1,000 uncorrelated structures, which were subsequently compared against each other.

The new Monte Carlo method reduced the number of potential energy calculations from ~500 per uncorrelated structure to ~40. This did not lead directly to a factor of 12 speed up for the simulation, although the time savings was quite substantial. The major reason for the nonlinear speed up was due to the slight decline in the acceptance ration and smaller step sizes were implemented. Figure 2.10 is the time correlation plot comparing the molecular translation simulation versus the new method.

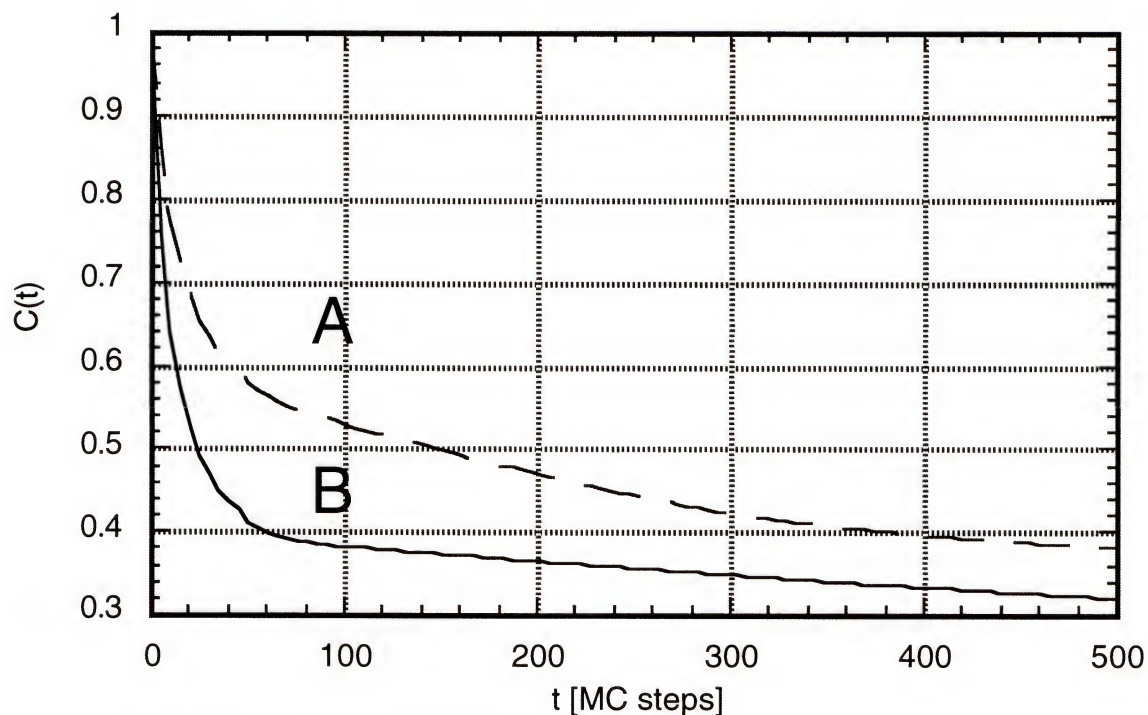


Figure 2.10: Monte Carlo Time Correlation Plot

- A Molecular Rotations and Translations
- B Molecular Rotations and Translations Coupled with Atomic Translations

The difference in the average energy for 100 uncorrelated structures generated from each of these methods is only 1.8 kcal/mol, while the radial distribution ($\text{O}\cdots\text{O}$) for a cluster of water molecules is statistically equivalent, Figure 2.11. The differences between the two Monte Carlo methods is small, although the modified version (A) has somewhat less pronounced radial distribution peaks. Since the two Monte Carlo methods are statistically equivalent, the modified Monte Carlo method described is better suited for hybrid QM/MM simulations, since it is $\sim 10\times$ faster.

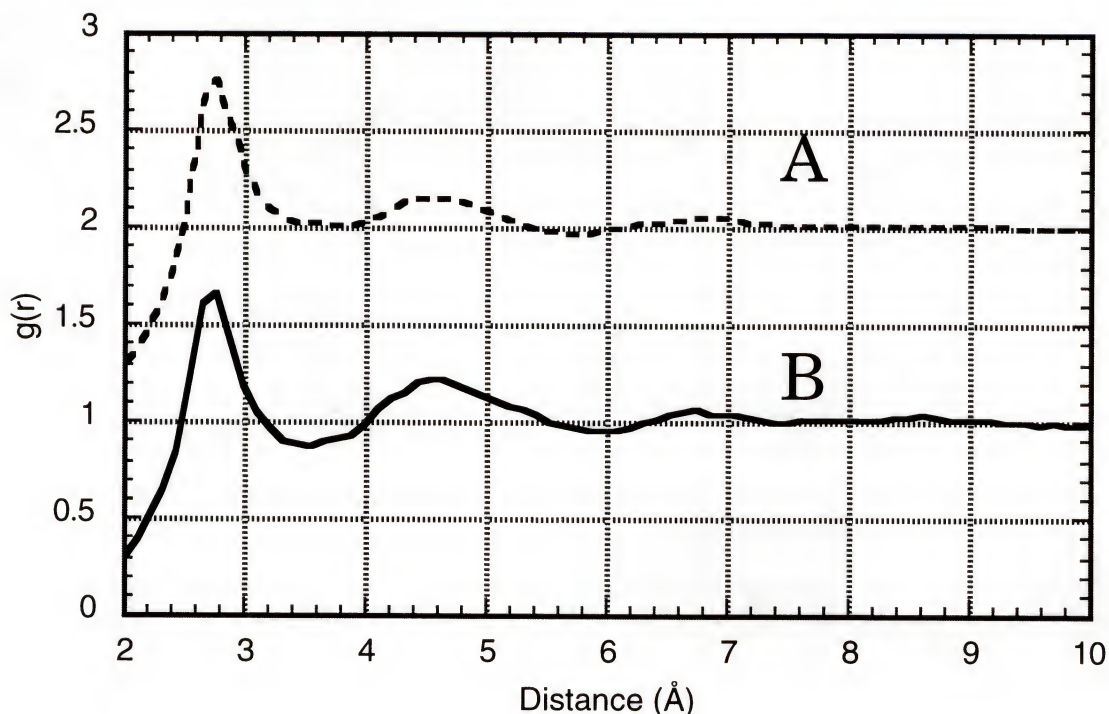


Figure 2.11: Monte Carlo Radial Distribution Function (O...O)

- A Molecular Rotations and Translations; 1 has been added to all of the $g(r)$ values for this line.
- B Molecular Rotations and Translations Coupled with Atomic Translations

Boundary Conditions

The current methods used for calculating the potential energy for a cluster of molecules is presently limited to significantly less than a million particles, while this initially sounds like a lot of molecules it is a minuscule fraction of Avogadro's number (6.022×10^{23}) or a mole of particles. Since current methods are incapable of modeling the entire system, a smaller model system is used for representing the "bulk" solution. An intrinsic

problem with modeling a smaller system is that the ratio of the surface area to the volume is very large. If this is uncorrected the system being modeled is a small drop instead of the solution in a flask. There are several methods that have been developed for embedding this drop in the remainder of the solution and reducing the edge effects. The various methods for correcting for the boundary conditions can be divided into two classifications. The first group is responsible for removing the edge effects and to make the model system resemble the interior region of a large system. These types of boundary conditions treat the boundary using explicit interactions and are therefore limited in size. The second classification is responsible for modeling the electrostatics of the bulk solution since the potential energy methods described previously are unable to calculate much more than a sphere of some 15 - 20 Ångstroms. The two different types of boundary conditions applied simultaneously provide an even more accurate model for describing long range interactions.

Periodic Boundary Conditions

Periodic boundary conditions are primarily used for removing the edge effects and reducing the effects arising from a large surface area to volume ratio. The periodic boundary conditions take the system of interest and create several images of this system displaced in such a way that the opposing edges are lined up [159]. This can be seen easily using a two dimensional square system (Figure 2.12). This method also incorporates a method for allowing molecules to enter and leave the inner core region. The movement of molecules between the images and core region is achieved by tracking the center of mass for each molecule. Then when the center of

mass for a specified molecule is moved outside the core region during the simulation, its image is moved into the core region from the opposite side.

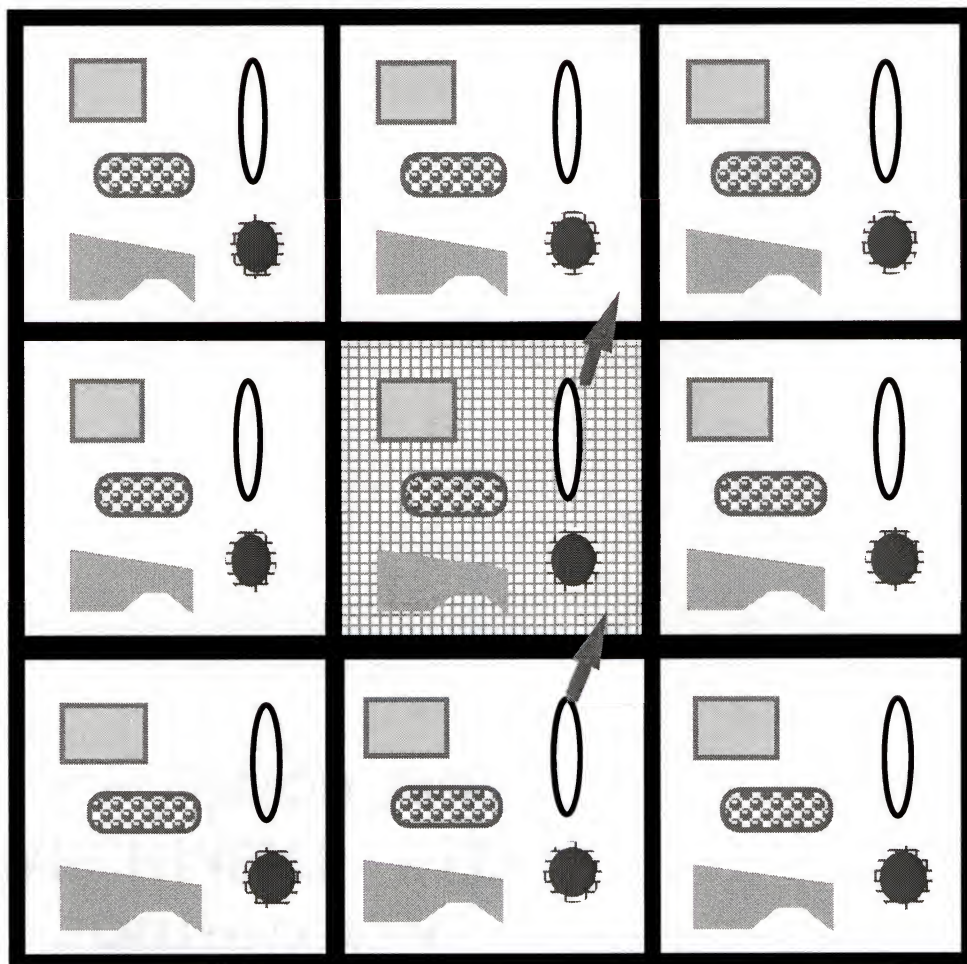


Figure 2.12: 2-D Square Periodic Boundary Conditions.

This 2-D square boundary conditions shows how a molecule is replaced when it leaves the simulation area.

This allows most statistical mechanics ensembles to be modeled. If the size of the core region remains constant, the number of particles in the core region will also remain constant, thus satisfying the constant particle

requirement for the canonical ensemble (NVT), microcanonical ensemble (NVE) and the isothermic-isobaric ensemble (NPT). This method can also accommodate ensembles that allow the number of molecules to change by allowing the number of molecules in the core region to change, such as the grand canonical ensemble (V, T, μ).

Stochastic Boundary Conditions

The stochastic boundary procedure requires that the system being studied can be partitioned into several regions. These regions are chosen based upon their spatial arrangement with respect to the primary volume of interest. The entire system is divided into a reaction region and a reservoir region. The reaction region contains the portion of the system that is of interest, such as the solute or the reaction center of a molecule. The reservoir region contains the remainder of the system that does not participate directly. While this type of boundary conditions is commonly used in conjunction with stochastic dynamics simulations, such as Brownian dynamics or Langevin dynamics, the method has also been used with molecular dynamics to study simple fluids and biomolecules [160]

The stochastic boundary conditions have several advantages over the periodic boundary conditions. First, the method does not need to simulate the image molecules. Secondly, the stochastic boundary conditions use a sphere for surrounding the simulated system, which provides a natural boundary for placing a dielectric continuum or some other field to represent the bulk solution.

Ewald Summations and Fast Multipole Methods

In recent years there has been numerous discussions pertaining to how to properly terminate long range Coulomb forces and energies [161-163]. It has been shown that assuming that the Coulombic interaction between two point charges is negligible after 10 - 15 Å is not accurate. Therefore, some method for accounting for the long range Coulombic interactions is required to properly model these systems. These terms are usually approximated using an Ewald summation [164,165] or fast multipole expansion [166-169].

The Ewald sum is a method for estimating the interactions between a molecule or ion and all of its periodic images (Figure 2.13). This method was originally developed by Madelung and Ewald to study ionic crystals [164,165].

The periodic boundary conditions are a severely truncated form of an Ewald sum. The potential energy for an Ewald sum is given by

$$V^z = \frac{1}{2} \sum_n \left(\sum_{i=1}^N \sum_{j=1}^N z_i z_j |\mathbf{r}_{ij} + \mathbf{n}|^{-1} \right) \quad 2.55$$

where z_i, z_j , are the point charges of the two interacting molecules or ions; the sum over n is the sum over all the simple cubic lattice points. The higher the sum index n , the more accurate and expensive the sum. There has been considerable work done to find approximate methods for

summing these interactions along with faster algorithms for calculating these interactions, such as the particle mesh Ewald sum method [170,171].

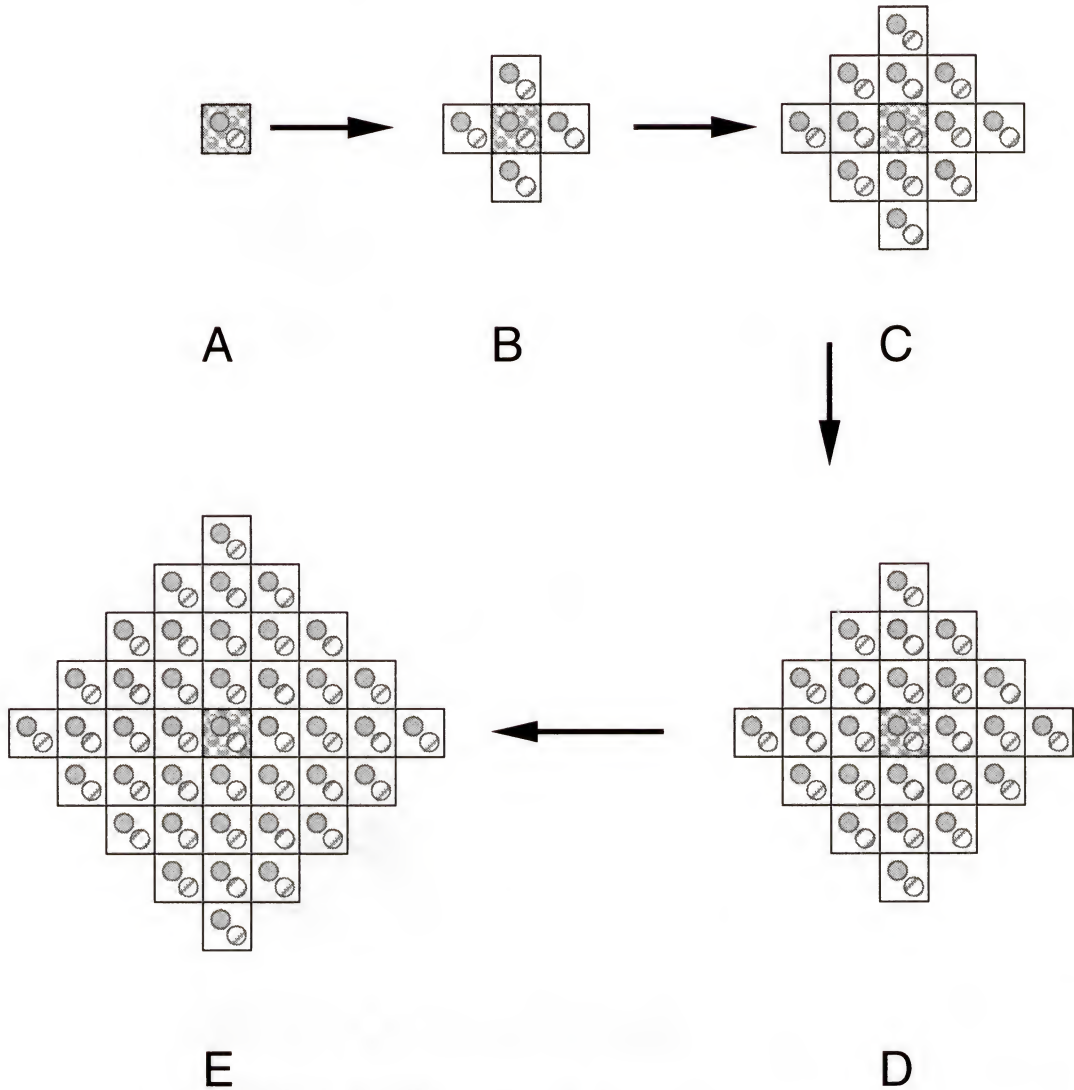


Figure 2.13: Building up of an Ewald Sum

A:	Simulation cell only	B:	1st Shell
C:	2nd Shell	D:	3rd Shell
E:	4th Shell		

The fast multipole expansion methods (FMM) are based upon a nonadaptive clustering scheme where every first level cell is subdivided an equal number of times to create new subcells. The number of cells grow exponentially with the refinement level of this method. This approach is best suited for a reasonably uniform distribution of particles, such as ionic crystal lattices. Using this type of method on a non uniform distribution has been achieved by using an adaptive form of the FMM [172], which is better suited for modeling solvents. In the adaptive scheme only the cells with a sufficient number of particles are subdivided. The adaptive scheme is directly related to the local particle density of the system.

While both of these methods have been used for describing solvent effects the periodic nature introduces some degree of artificial periodicity into the solvent. Therefore I have decided to also examine using a dielectric continuum for modeling the long range interactions.

Self-Consistent Reaction Fields

The use of a continuum model for representing the solvent is based upon the classical electrostatic treatment of interacting systems by Born, Onsager, and Kirkwood [173-175]. The total energy of a molecule or ion with a defined geometry and conformation can be written as

$$E_{total} = E_{solute} + E_{solvation} \quad 2.56$$

with E_{solute} being the energy found in an isolated gas phase molecule.

The solvation energy, $E_{solvation}$, is made up of three components,

$$E_{\text{solvation}} = E_{\text{es}} + E_{\text{dis}} + E_{\text{cav}} \quad 2.57$$

E_{es} is the electrostatic solvent-solute binding energy arising from the interaction of permanent and induced electric moments, E_{dis} is the interaction energy resulting from dispersion forces, and E_{cav} is the energy required to form a cavity in the solvent to accommodate the solute molecule. The first two terms in Equation 2.57 are negative while the last term is always positive. Each of these terms may depend parametrically on quantities that vary as a function of the molecular conformation or geometry.

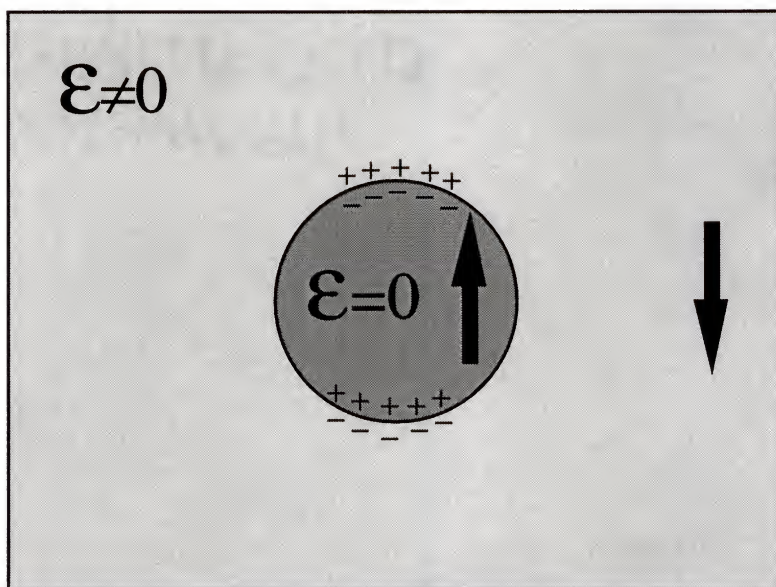


Figure 2.14: Self Consistent Reaction Field

ϵ is the macroscopic dielectric permittivity of the continuum solvent.

The calculation of the solute energy (E_{es}) is treated as a point dipolar ion of charge Q and the total multipole moment expansion (m) at the center of a sphere of effective radius (α) imbedded in the solvent. The solvent is represented as a polarizable dielectric continuum of dielectric constant ϵ . The solute induces a reaction field (E_R) in the solvent that then acts back on the solute, (Figure 2.14).

If the species is charged, $E_{es}^{(0)}$ is given by the famous Born term [173],

$$E_{es}^{(0)} = -\frac{1}{2} Q^2 \left(1 - \frac{1}{\epsilon}\right) \frac{1}{\alpha^3} \quad 2.58a$$

while non-charge species (multipole moment expansion) are given by,

$$E_{es}^{(0)} = -\frac{1}{2} m * E_R \quad 2.58b$$

where the dielectric continuum energy is given by,

$$E_R = \left(1 - \frac{1}{\epsilon}\right) \frac{m}{\alpha^3} \quad 2.59$$

Both m and α depend on conformation; m is computed from the wave function and α is the molecular cavity radius defined by equating the molecular volume of a sphere (V),

$$\alpha = \left(\frac{3}{4\pi} V \right)^{1/3} \quad 2.60$$

Where the molecular volume (V) is estimated from the Cartesian coordinates of the atoms of the solute.

The dispersion energy (E_{dis}) can be estimated using

$$E_{dis} = \frac{\rho}{2} \int_0^{\infty} v^{eff}(r) g^{(2)}(r) 4\pi r^2 dr \quad 2.61$$

where ρ is the number density of the solvent; v^{eff} is the effective pairwise potential function for the solute-solvent interaction and $g^{(2)}$ is a radial distribution function. For simplicity $g^{(2)}(r)$ is taken to be zero for $r < \alpha$ and unity for $r > \alpha$. The dispersion energy term can also be determined from the experimentally known fact that larger molecules have a higher dispersion energy, and the dispersion energy is roughly proportional to the volume of the molecule:

$$E_{dis} = \zeta V \quad 2.62$$

where V is the molecular volume, and ζ is a proportionality factor.

The cavitation energy (E_{cav}) is estimated from the cavity surface area (A) and the solvent surface tension (γ) as

$$E_{cav} = f 4\pi \alpha^2 \gamma \quad 2.63$$

where f is a factor that relates macroscopic dimensions to microscopic dimensions.

While the previous derivation of the dielectric continuum model is useful for understanding the various types of interactions involved, I prefer the simpler Fock operator derivation. The self consistent reaction field, restricted to the dipole term for simplicity, is given by a modified Fock equation,

$$F' = F_o - \vec{\mu} \bullet \vec{g} \bullet \langle \Phi | \mu | \Phi \rangle \quad 2.64$$

where F_o is the Fock operator for the gas phase calculation; μ is the dipole moment of the continuum solvent; Φ is the wavefunction describing the isolated molecule; and g is the Onsager term given by

$$g = \frac{(\epsilon + 1)}{(2\epsilon - 1)} \alpha^{-3} \quad 2.65$$

at the SCF level of theory, where ϵ is the macroscopic dielectric permittivity of the continuum solvent; α is the cavity radius needed for the solute.

The development of how to model an absorption of a photon light as an instantaneous process for the configuration interaction (CI) theory is presented in reference [53]. The difference between modeling a solvent

relaxed ground state interaction and the instantaneous light absorption method used in the CI SCRF calculations are shown in Figure 2.15. The CI SCRF theories have the additional approximation that only the electron polarization of the solvent is allowed to react to the light absorption.

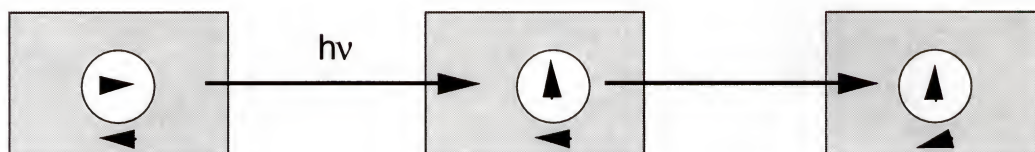


Figure 2.15: Comparison of the Reaction Field for Ground State and Excited States.

The first box is the ground state solute and relaxed reaction field. The second box is the excited solute and the relaxed ground state reaction field. The third box is the excited state solute with the electronically relaxed reaction field, solvent electron relaxation only.

The dielectric continuum model methods have demonstrated that they are capable of modeling certain types of solvent interactions [54]. These methods require additional considerations when the solvent is capable of hydrogen bonding to the solute [55-57].

CHAPTER 3 TEST SYSTEMS

Introduction

The objective of this chapter is to demonstrate that the various methods used for investigating large molecular systems have been correctly implemented. The testing of these various techniques has been accomplished using small “ideal” systems for which either experimental or other theoretical results are available for comparison. The testing of the different methods is essential for two reasons: first, to see if the theory has been coded correctly; second, to determine the limitations of these theories. The first section of this chapter pertains to the testing of the molecular mechanics algorithm and the stochastic methods that have been written for use in the QM/MM implementation. The last part of this chapter involves testing of the hybrid quantum mechanics-molecular mechanics (QM/MM) method. The QM/MM section is further divided into two sections, the first section involves testing the QM/MM versus other QM/MM methods, *ab initio* calculations and experimental data. The second section investigates the consistency of the QM/MM method when studying explicit solvent interactions.

Molecular Mechanics and Stochastic Methods

There are several different molecular mechanics force fields that have been implemented into the ZINDO program package [65]. The SPC

[75] force field has been chosen for modeling water while the AMBER [72] and the consistent force field (CFF) [68] model the remainder of the interactions. The combination of these force fields is restricted to organic compounds but additional Lennard-Jones (12-6) parameters have been added for some of the transition metals. The molecular mechanics method was tested by comparing the results obtained with this implementation with the results obtained using a commercial software package using the same force field.

Testing of the molecular dynamics and Monte Carlo methods was accomplished by running molecular mechanics simulations of 99 water molecules using both methods and comparing the results to one another and to published results [18,176]. The testing for the stochastic methods was accomplished by comparing the oxygen-oxygen and oxygen-hydrogen radial distribution functions.

The molecular mechanics and stochastic techniques have been added to the ZINDO CVS repository. The CVS repository code is subjected to testing under three different UNIX operating systems every time that code is modified. Using CVS for maintaining the source code has several advantages for preventing the addition of new bugs into the code. The first advantage is that the test cases are run every time the code changes, therefore the code must perform the same before and after any changes. Additionally this method has the advantage of reproducing the source code from a specified date for direct comparison against the new version. Finally this method produces a log of the changes made along with who made the changes. The use of CVS ensures that the tested source code remains functional and reduces the possibility of introducing new bugs into the existing code.

Hybrid Quantum Mechanics-Molecular Mechanics

Objective

The hybrid quantum mechanics-molecular mechanics (QM/MM) method, implemented into the ZINDO program package, has been tested against the QM/MM method described by Field et al [120]. For the remainder of this chapter the QM/MM method implemented in the ZINDO program package will be referred to as INDO/MM and the work done by Field as AM1/MM. The first three examples in this section compare high level *ab initio* calculations and experimental results to the INDO/MM and AM1/MM methods. The remaining examples are used to test the types of interactions that the INDO/MM method can consistently model well. These latter test cases are modeling hydrogen bonding interactions that are frequently encountered in organic and biological chemistry.

One of the more important features of a QM/M procedure is to give the same results regardless of the selection of the regions. This is in practice not true, since the QM region is needed to model phenomena that the MM region cannot. The following examples are used to measure how consistent the QM/MM method is at modeling the gas phase dimer (X/Y) test cases, to determine the size of the errors associated with the selection of the method used for modeling a molecule. This is accomplished by modeling X as the QM and Y as the MM (QM/MM) model and modeling X as the MM and Y as the QM (MM/QM). Ideally these two models would be equivalent, indicating that the model describing the interaction between the QM and MM regions is balanced.

Water Dimer

Hydrogen bonding plays a critical role in many chemical reactions both in solution and in biological processes. The water dimer is the simplest model for hydrogen-bonding phenomenon. It has been investigated in great detail, using classical force fields [177,178], semi-empirical methods [179,180], *ab initio* methods including electron correlation [181-185], and a hybrid QM/MM model [120]. At least three different stable conformers have been identified [120,177-179,181-187]. The most stable conformer has C_s symmetry with an almost linear hydrogen bond. A second low energy conformer possesses C_{2v} symmetry and has two hydrogen bonds, while a third conformer has almost C_{2h} symmetry and possesses three hydrogen bonds. The three lowest energy conformers of the water dimer are displayed in Figure 3.1 - 3.3.

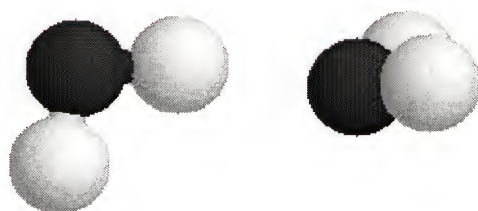


Figure 3.1: Water Dimer C_s Symmetry Conformation

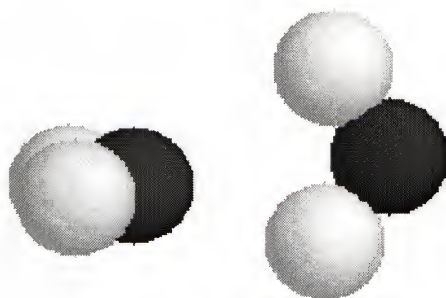


Figure 3.2: Water Dimer C_{2v} Symmetry Conformation

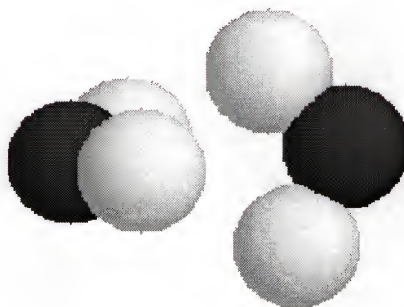


Figure 3.3: Water Dimer Cyclic Conformation

The first comparison of the QM/MM model is to test and see if a totally symmetric dimer (Figure 3.3) is independent of the model used to represent either water molecule. Taking the symmetric cyclic conformations and optimizing the conformer while constraining the symmetry showed that this implementation of the QM/MM procedure is independent upon the selection of the QM and MM regions for totally symmetric identical molecules, Table 3.1. The indistinguishability between

the two water molecules is lost when the water molecules are no longer identical as seen from the other conformations.

Table 3.1: Cyclic Water Dimer Results

	Distance H-O...H (Å)	Distance O...O (Å)	Angle (degree)	ΔE (kcal/mol)
INDO	2.27	2.68	xxx	3.8
INDO/MM	2.6387	3.0866	36.05	0.6397
INDO/MM	2.6387	3.0866	36.05	0.6397

Angle is H(1) - O(2) - H(1), where 1 and 2 represent the molecule number

The geometries and hydrogen-bonding energies of the water dimers in the C_s conformation as predicted by the INDO model and the INDO/MM model are collected in Table 3.2, along with data from experimental observations and other theoretical predictions. Some care has to be taken when comparing experimental observations, theoretical geometries, and bonding energies, since theoretical calculations give bond energies without inclusion of the zero-point vibration energy. Furthermore, the size of the basis set, electron correlation, and basis set superposition error (BSSE) can have a large effect on the predicted geometries and bond energies for quantum mechanics *ab initio* calculations reported. These effects on the predicted geometry and binding energy of the water dimer were investigated in detail by Frisch et al. [181]. The theoretical predicted geometry with a medium-sized basis set, such as 6-31G(d), at the second-order perturbation level (MP2) of theory reproduces the observed experimental geometry well. In a recent detailed study of the water dimer,

Feyereisen et al. [185] estimated the binding energy to be 5.0 ± 0.1 kcal/mol. In the study by Feyereisen, very large basis sets were used (up to 547 functions) and BSSE was carefully examined.

Table 3.2: Summary of the Water Dimer, As Predicted at Different Levels of Theory

	Distance H-O...H (Å)	Distance O...O (Å)	Angle (degree)	ΔE (kcal/mol)
INDO C_s	1.31	2.33	178.0	26.2
SPC C_s	1.74	2.75	154.0	6.6
AM1/MM C_s	2.02	---	159.5	3.3
MM/AM1 C_s	1.99	---	169.0	4.6
INDO/MM C_s	1.89	2.87	175.4	3.3
MM/INDO C_s	1.87	2.87	176.4	3.0
ab initio ^{a,b} C_s	1.95	2.91	175.5	5.0 ± 0.1
observed ^{c,d}	---	2.98	---	5.4 ± 0.5

a MP2(FC)/6-311++G(2d,2p) geometry from ref. 181

b M.W. Feyereisen, D. Feller, and D.A. Dixon ref. 185

c T.R. Dyke, K.M. Mack, and J.S. Muenter ref. 188

d L.A. Curtiss, D.J. Frurip, and M.J. Blander ref. 189

Clearly the INDO model does not reproduce the geometry or the binding energy of the water dimer, as seen in Table 3.2. The molecular mechanics SPC water model does not reproduce the geometry well either, although the binding energy is reproduced reasonably well [190]. The proposed INDO/MM model significantly improves the description of the geometry when compared with either the INDO or MM methods. The

binding energy is also significantly improved over the INDO model. The hybrid QM/MM Hamiltonian developed in chapter 2 (INDO/MM) produces a more consistent description of the water dimer than does the AM1/MM hybrid Hamiltonian proposed by Field et al. [120], as demonstrated by comparison of the QM/MM and MM/QM results for each QM model. The results obtained from the C_s conformation were typical of those obtained for any conformation that was not totally symmetric.

Formamide/Water Complex

The formamide water complex has been studied with theoretical methods by many groups [181-185]. One of the reasons for the large interest in this complex is that formamide can be considered the simplest molecule containing a peptide bond. Four conformations have been identified, all possessing C_s symmetry [191,192]. The four minimum-energy structures can be seen in Figure 3.4 - 3.7.

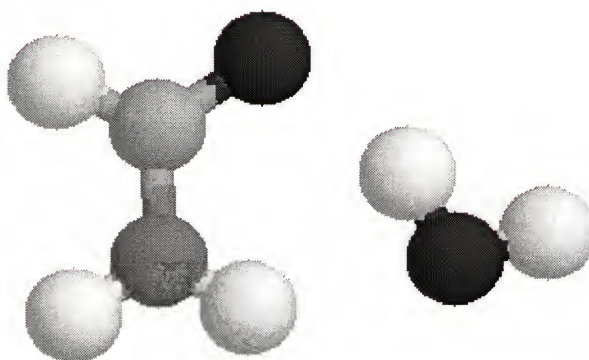


Figure 3.4: Water/Formamide Complex; Conformer 1

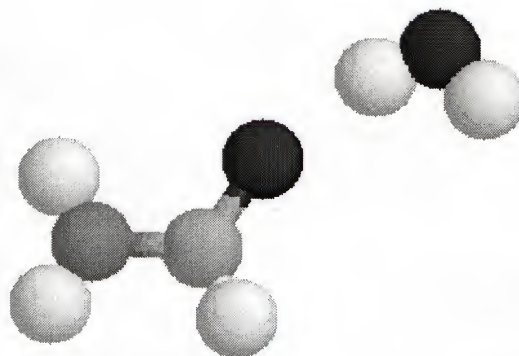


Figure 3.5: Water/Formamide Complex; Conformer 2

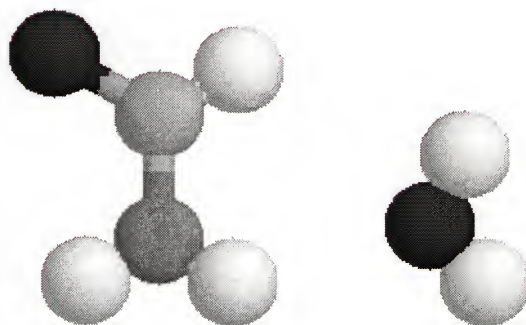


Figure 3.6: Water/Formamide Complex; Conformer 3

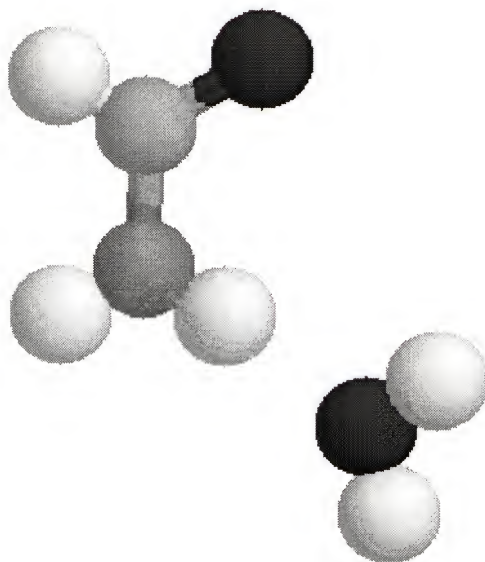


Figure 3.7: Water/Formamide Complex; Conformer 4

The formamide can act as both a proton acceptor and proton donor in hydrogen bonding. The two hydrogen bonds formed with water are of almost equal binding strength. However, the carbonyl oxygen binding site (conformer 2) is slightly favored energetically over the amide hydrogen binding site (conformers 3 and 4). Conformer 1, which has two hydrogen bonds in a cyclic arrangement, was found to be the global energy minimum within the C_s symmetric potential surface.

The *ab initio* potential energy surface was found to be very flat, and some of the conformers in Figures 3.4 - 3.7 indicate that large changes in bond angles induce only very small energy changes. Some nonsymmetric points on the full potential energy surface were investigated by Jaisen and Stevens [191]. They concluded that it was likely that the global energy minimum corresponds to a nonsymmetric geometry. Some additional

MP2/cc-pVDZ calculations of the complex have been performed in order to estimate the importance of electron correlation on the geometry and the binding energy. One calculation of the MP2/cc-pVDZ was done with C_s symmetry constraints starting from a geometry close to conformer 1. A second geometry optimization was performed without symmetry constraints, starting from the INDO/MM optimized geometry. These calculations suggest that the C_1 symmetric MP2/cc-pVDZ geometry is favored by 1.3 kcal/mol over conformer 1 (C_s). The symmetry breaking leads to a complex that is still cyclic and possesses two hydrogen bonds, but the water molecule has been rotated so that the plane of the water is perpendicular to that of the formamide. However, the binding energy after BSSE correction is smaller for the C_1 symmetric conformer compared to the C_s symmetric conformer.

The INDO/MM binding energies and hydrogen bond geometries are summarized in Tables 3.3 - 3.6 for comparison. All the INDO/MM optimized geometries are done without symmetry constraints. It is possible to optimize the structures of conformers 1 and 2 while preserving a plane of symmetry (C_s) by starting with a planar initial geometry (The current version of INDO/MM does not take advantage of spatial symmetry). The energies of the complexes of C_s symmetry differ very little from the corresponding ones of C_1 symmetry. The INDO/MM geometries and binding energies are again more consistent than those obtained from the AM1/MM model [120].

Table 3.3: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 1

	<i>ab</i> <i>initio</i>	MP2 C _s	MP2 C ₁	AM1/ MM	MM/ AM1	INDO/ MM	MM/ INDO
ΔE	7.8	7.2	6.8	6.7	5.0	7.5	6.4
	---	(11.8)	(13.1)	---	---	---	---
r1	2.16	1.96	1.98	2.45	2.07	2.05	2.19
r2	2.06	1.94	1.93	2.03	2.55	1.84	1.80
A1	83.7	84.5	79.2	67.3	84.9	78.8	68.4
A2	138.6	139.0	141.3	136.4	160.6	134.5	132.4
A3	110.1	105.6	104.9	112.5	112.3	107.9	107.4
A4	143.3	146.7	151.4	161.3	126.9	149.8	162.8

ΔE in kcal/mol

r1 O(W) ... H(F)

A1 H(W) - O(W) ... H(F)

A3 C(F) - O(F) ... H(W)

ab initio results from 191

r2 O(F) ... H(W)

A2 O(W) ... H(F) - N(F)

A4 O(F) ... H(W) - O(W)

Numbers in Parenthesis are the Binding Energies before BSSE correction.

Table 3.4: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 2

	<i>ab initio</i>	AM1/MM	MM/AM1	INDO/MM	MM/INDO
ΔE	5.6	5.7	2.9	5.4	3.3
r1	2.03	2.12	2.23	1.62	1.70
A1	118.8	124.4	141.4	112.0	107.9
A2	169.0	175.6	150.0	175.4	174.4

ΔE in kcal/mol

A1 C(F) - O(F) ... H(W)

r1 O(F) ... H(W)

A2 O(F) ... H(W) - O(W)

Table 3.5: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 3

	<i>ab initio</i>	AM1/MM	MM/AM1	INDO/MM	MM/INDO
ΔE	5.1	4.0	5.2	4.3	3.6
r1	2.12	2.01	1.99	1.92	1.95
A1	175.9	182.9	176.1	176.2	169.3
A2	178.5	182.7	173.3	---	---

ΔE in kcal/mol

A1 N(F) - H(F) ... O(W)

r1 O(W) ... H(F)

A2 H(F) ... O(W) - H(W)

Table 3.6: Binding Energies and Geometries for the Formamide/Water (F/W) Complex, Conformer 4

	<i>ab initio</i>	AM1/MM	MM/AM1	INDO/MM	MM/INDO
ΔE	5.1	5.2	5.2	4.7	3.3
r1	2.11	2.46	2.03	2.00	1.99
A1	177.1	137.8	164.1	144.9	134.9
A2	121.5	82.7	109.6	---	---

ΔE in kcal/mol

A1 H(W) - O(W) ... H(F)

r1 O(W) ... H(F)

A2 O(W) ... H(F) - N(F)

The INDO/MM results compare well with the *ab initio* results and the relative ordering of the stability for the different conformers is reproduced, in contrast to the AM1/MM model. The largest geometry difference between the *ab initio* results and the INDO/MM results is observed when the carbonyl group is a proton acceptor for the hydrogen bond. The O...H bond is calculated to be too short in all cases by the INDO/MM model. It is likely that this deficiency could be corrected by modifying the Lennard-Jones parameter for the carbonyl oxygen. However, at the present stage I do not want to do a complete reparametrization for every atom type, but rather only use a linear scaling of the MM-MM Lennard-Jones parameters. Furthermore, the MP2 predicted geometry of conformer 1 has much shorter hydrogen bond lengths than the Hartree-Fock SCF predicted geometry, and it is the experimental results that our model is trying to reproduce.

Formamide/Methanol Complex

The formamide/methanol complex was also studied by Jaisen and Stevens [191]. They reported four low-energy conformations, similar to those found for the formamide/water complex, see Figure 3.8 - 3.11.

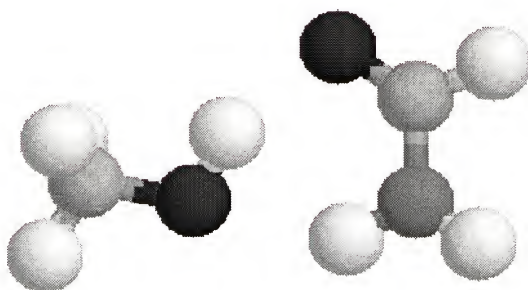


Figure 3.8: Methanol/Formamide Complex; Conformer 1

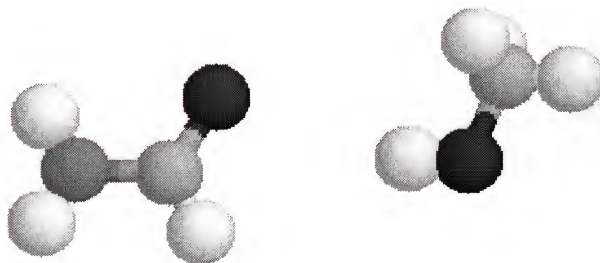


Figure 3.9: Methanol/Formamide Complex; Conformer 2

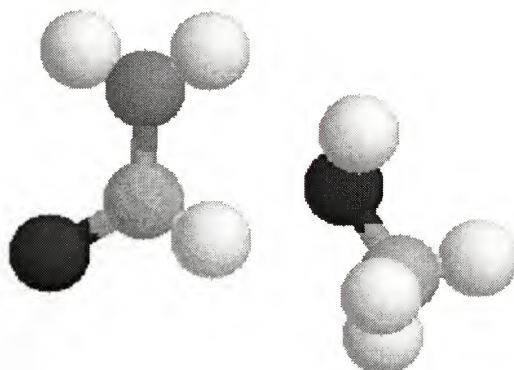


Figure 3.10: Methanol/Formamide Complex; Conformer 3

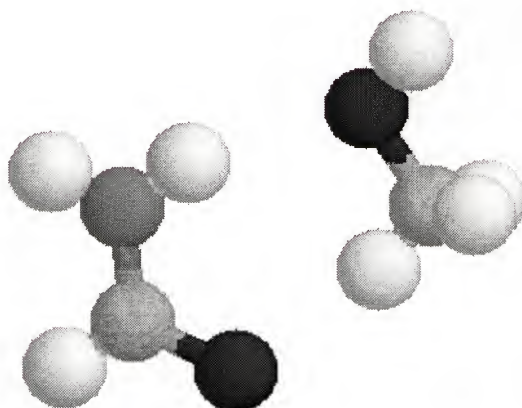


Figure 3.11: Methanol/Formamide Complex; Conformer 4

The cyclic double-hydrogen bonded conformer 1 is predicted to be the most stable, as it was in the formamide/water complex. The *ab initio* predicted ordering of the relative stability for the four conformers is not reproduced by either of the hybrid QM/MM models, but both hybrid models predict the cyclic conformer as the most stable one; see Tables 3.7 - 3.10.

Table 3.7: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 1

	HF/DZ	HF/DZP	AM1/MM	MM/AM1	INDO/MM	MM/INDO
ΔE	7.6	7.8	6.4	6.1	9.2	11.4
r1	2.08	2.17	2.32	1.97	2.17	2.25
r2	1.94	2.05	2.17	2.42	2.08	2.10
A1	85.5	83.1	---	---	89.6	58.8
A2	139.3	139.3	139.3	153.7	140.9	122.0
A3	109.8	110.0	115.0	109.9	113.9	95.7
A4	144.1	144.3	139.3	128.2	134.3	136.0

ΔE in kcal/mol

r1 O(M) ... H(F)

A1 H(M) - O(M) ... H(F)

A3 C(F) - O(F) ... H(M)

ab initio results from 191

r2 O(F) ... H(M)

A2 O(M) ... H(F) - N(F)

A4 O(F) ... H(M) - O(M)

Table 3.8: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 2

	<i>ab initio</i>	AM1/MM	MM/AM1	INDO/MM	MM/INDO
ΔE	5.6	5.6	2.9	4.7	6.9
r1	1.93	2.13	2.32	2.06	2.02
A1	173.0	168.3	150.5	174.9	175.0
A2	136.3	141.6	97.5	138.1	107.3

 ΔE in kcal/mol

A1 C(F) - O(F) ... H(M)

ab initio results from 191

r1 O(F) ... H(M)

A2 O(F) ... H(M) - O(M)

Table 3.9: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 3

	<i>ab initio</i>	AM1/MM	MM/AM1	INDO/MM	MM/INDO
ΔE	4.5	3.1	5.6	6.2	4.9
r1	2.00	1.98	1.91	2.12	2.14
A1	174.5	169.9	177.3	141.2	148.1
A2	127.3	156.8	125.6	92.4	94.6

 ΔE in kcal/mol

A1 N(F) - H(F) ... O(M)

ab initio results from 191

r1 O(M) ... H(F)

A2 H(F) ... O(M) - C(M)

Table 3.10: Binding Energies and Geometries for the Formamide/Methanol (F/M) Complex, Conformer 4

	<i>ab initio</i>	AM1/MM	MM/AM1	INDO/MM	MM/INDO
ΔE	4.7	2.8	5.8	5.5	7.8
r1	1.95	2.02	1.93	2.16	1.95
A1	176.6	162.1	174.5	148.9	175.5
A2	118.4	91.3	108.5	83.2	92.2

ΔE in kcal/mol

A1 C(M) - O(M) ... H(F)

ab initio results from 191

r1 O(M) ... H(F))

A2 O(M) ... H(F) - N(F)

Once again, we can see that the INDO/MM model produces more consistent results than does the AM1/MM model [120]. The INDO/MM predicted binding energies are in most cases larger than the *ab initio* predicted binding energies. The INDO/MM predicted hydrogen bond lengths are generally too short compared to the *ab initio* predicted values for the formamide/water complex and too long for the formamide/methanol complex. However, inclusion of electron correlation and use of larger basis sets change the *ab initio* predicted geometries greatly, so a definitive conclusion is hard to draw from the above comparison.

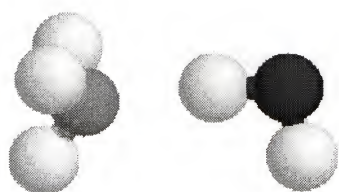
Solvent Dimer Testing

One of the major advantages of a QM/MM method over the dielectric continuum models is that it is capable of explicitly modeling solute-solvent interactions. While these specific interactions are not crucial when modeling most organic solvents, they are important when strong interactions between the solute and solvent are present, such as a hydrogen bond between the solute and solvent. The explicit solute-solvent interactions, therefore, must be examined in some detail to ensure that the QM/MM method adequately represents these interactions. Several of these hydrogen bonding interactions have been examined using water and ammonia coupled to several small organic hydrogen bonding compounds. The QM/MM method is compared to the PM3 semi-empirical method to insure that the QM/MM method is not only consistent with itself but is in agreement with other theoretical methods. The PM3 semi-empirical method was chosen as the reference since it has been shown to reproduce hydrogen bonding complexes [193].

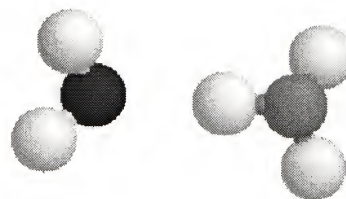
Water / Ammonia

Modeling the water-ammonia complex is important for determining if this method is capable of modeling inorganic compounds that contain ammonia ligands. This type of interaction is also important for determining how well this method can reproduce water/amino group interactions found in organic and biological compounds. The water-ammonia complex has two possible hydrogen bonding conformers that have been examined, (Figure 3.12). The first conformer has the water molecule as the hydrogen bond donor, while the second conformer has the

ammonia molecule as the hydrogen bond donor. The binding energy and hydrogen bonding distances and angle are given in Table 3.11 for the first conformer and in Table 3.12, for the second conformer.



conformer 1



conformer 2

Figure 3.12: Water / Ammonia Complex

Table 3.11: Water / Ammonia (W/A) Dimer Results; Conformer 1

	INDO	MM	PM3	INDO/MM	MM/INDO
ΔE	18.19	10.82	3.04	3.86	3.21
R1	1.459	1.758	1.825	1.812	1.791
R2	2.481	2.749	2.793	2.786	2.764
A1	178.0	159.9	179.3	160.6	164.5

All Energies in kcal/mol
 R1 = N(A) ... H(W) R2 = N(A) ... O(W)
 A1 = N(A) ... H(W) - O(W)

Table 3.12: Ammonia / Water (A/W) Dimer Results; Conformer 2

	INDO	MM	PM3	INDO/MM	MM/INDO
ΔE	17.84	6.44	-0.88	1.78	2.19
R1	1.362	1.718	1.866	1.908	1.950
R2	2.441	2.748	2.866	3.008	2.964
A1	173.1	173.1	170.9	175.6	176.1

All Energies in kcal/mol
 R1 = O(W) \cdots H(A) R2 = N(A) \cdots O(W)
 A1 = N(A) - H(A) \cdots O(W)

The geometries obtained from the INDO/MM and MM/INDO models are in good agreement with each other. The first conformer has an average difference of ~ 0.02 Å for the two measured distances while the second conformer has an average error of 0.04 Å, these errors are of the same order of magnitude as those expected using semi-empirical methods. The difference in the measured angle is 4° for conformer 1 and 1° for conformer 2. This error is also of the same order of magnitude expected for a semi-empirical method. The average geometric differences between the QM/MM methods and the PM3 results are small also. This is a good indication that this method is correctly modeling the interactions. The binding energies predicted using the INDO/MM and MM/INDO models are also in good agreement, approximately 0.5 kcal/mol difference for both conformers. Comparing the PM3 predicted binding energies to the QM/MM models is not as good. The first conformer differs from the PM3 results by ~ 0.5

kcal/mol, while the second conformer has a difference of 2.5 kcal/mol. The PM3 calculations found a local minima in this configuration although this minima is not energetically allowed. Overall the results obtained using the QM/MM model are in good agreement with each other and the results obtained using the PM3 semi-empirical method.

Water/Methanol

The water / methanol complex is important for determining how well the QM/MM method can model alcohol/water interactions and can provide some insight into how this method can model the carboxylic acid / water interactions. Again, as in the water-ammonia case, there are two possible hydrogen bonding conformers that were examined, (Figure 3.13). The binding energy and hydrogen bonding information are given in Table 3.13 for the first conformer and in Table 3.14 for the second conformer.



Figure 3.13: Water/Methanol Complex

Table 3.13: Water/Methanol (W/M) Dimer Results; Conformer 1

	INDO	PM3	INDO/MM	MM/INDO
ΔE	21.64	2.95	3.23	3.24
R1	1.340	1.816	1.740	1.760
R2	2.364	2.774	2.742	2.752
A1	177.3	175.3	173.0	175.3

All Energies in kcal/mol
 R1 = O(M) ... H(W) R2 = O(M) ... O(W)
 A1 = O(M) ... H(W) - O(W)

Table 3.14: Methanol/Water (M/W) Dimer Results; Conformer 2

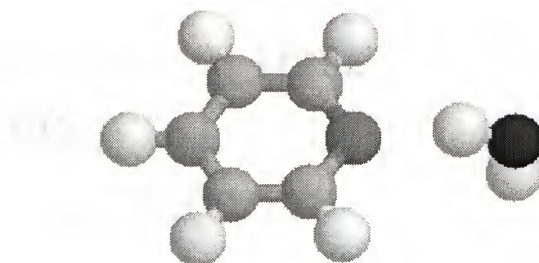
	INDO	PM3	INDO/MM	MM/INDO
ΔE	23.57	3.20	3.29	1.351
R1	1.323	1.814	1.980	1.926
R2	2.353	2.753	2.9604	2.877
A1	178.7	165.9	166.9	172.6

All Energies in kcal/mol
 R1 = O(W) ... H(M) R2 = O(M) ... O(W)
 A1 = O(M) - H(M) ... O(W)

The average distance difference when comparing the INDO/MM and MM/INDO models is ~ 0.02 Å for conformer 1 and ~ 0.07 Å for conformer 2. The difference in hydrogen bonding angle is $\sim 2^\circ$ for conformer 1 and 7° for conformer 2. The differences in the geometries for conformer 1 are satisfactory while those obtained from conformer 2 are at the upper limits of the acceptable range. The difference in the binding energy for conformer 1 is only 0.01 kcal/mol while the difference for conformer 2 is ~ 2 kcal/mol. The inconsistency of the two models is mostly due to the molecular mechanics force field. The force field used for this calculation did not have a binding minima for conformer 2. The size of this discrepancy could be reduced by switching to a better force field, such as the MMx force fields by Allinger [67,69]. While reparameterization of the QM/MM interface may increase the consistency of the two models, this would also reduce the generality of the method, since this reparameterization would have to be done for every force field.

Water/Pyridine

The water-pyridine interaction is of vital importance for modeling biological systems that we are interested in investigating. This interaction is present in biological systems such as porphyrins and derivatives of porphyrins. There is only one hydrogen bonding conformer that was investigated for this system (Figure 3.14). While there are several hydrogen bonding conformations available, this conformation was determined to be slightly favored energetically from simulated annealing techniques. The binding energy and hydrogen bonding geometric information is presented in Table 3.15.



conformer 1

Figure 3.14: Water/Pyridine Complex

Table 3.15: Water/Pyridine (W/P) Dimer Results

	INDO	PM3	INDO/MM	MM/INDO
ΔE	17.68	2.35	3.18	3.18
R1	1.442	1.844	1.791	1.836
R2	2.480	2.810	2.735	2.778
A1	174.0	177.6	154.1	156.6

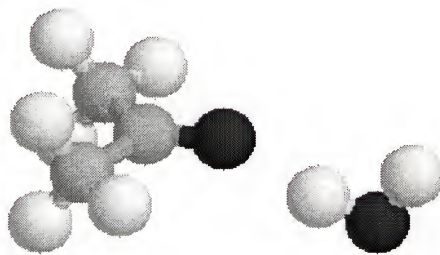
All Energies in kcal/mol
 R1 = N(P) ... H(W) R2 = N(P) ... O(W)
 A1 = N(P) ... H(W) - O(W)

The INDO/MM and MM/INDO method agree exceptionally well for the calculated binding energy, no difference to 5 significant figures. The two measured distances were also in agreement, approximately 0.04 Å difference in the two models. The QM/MM method also agreed with the PM3 predicted values for the bond distances, although the QM/MM method

was consistently shorter by ~ 0.05 Å. This shortening of the nitrogen-oxygen bond distance would also account for the smaller binding energy observed with PM3 when compared to the QM/MM method.

Water/Acetone

The interactions found in the water-acetone complex are representative of those observed in aldehydes, ketones, and carboxylic acids. Interactions of this type are also important for modeling amino acids and nucleotides, which are important molecules in biological systems. There is only one hydrogen bonding conformer examined for this complex (Figure 3.15). This conformer was chosen from simulated annealing simulations to have the strongest binding energy using molecular mechanics. The binding energy and hydrogen bonding information for this system are given in Table 3.16.



conformer 1

Figure 3.15: Water/Acetone Complex

Table 3.16: Water/Acetone (W/A) Dimer Results

	INDO	PM3	INDO/MM	MM/INDO
ΔE	25.91	2.39	6.53	3.98
R1	1.310	1.827	1.615	1.779
R2	2.336	2.774	2.608	2.758
A1	174.4	169.2	166.2	167.0

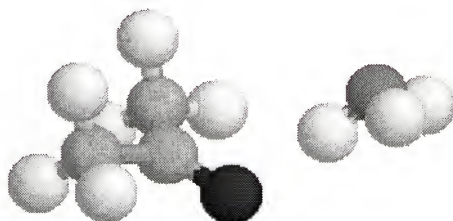
All Energies in kcal/mol
 R1 = O(A) ... H(W) R2 = O(A) ... O(W)
 A1 = O(A) ... H(W) - O(W)

The INDO/MM and the MM/INDO models failed to obtain the same minimum energy structure for this complex. The difference in binding energy for this system was ~ 3 kcal/mol, and the average difference in the measured distances was $\sim 0.2\text{\AA}$. While this error is quite large, the MM/INDO model is similar to the results obtained using PM3, and this is the system that has acetone as the quantum mechanics region. The two most probable sources of this inconsistency are the mismatch of the two force fields or that the CFF force field does not model acetone accurately. The mismatch in the quality of the force fields is a probable source of error in all of the dimers studied, since the SPC force field is used for representing water molecules while the CFF or AMBER force fields represent the remaining molecules. The force field error could also be associated with the determination of the point charges, since the SPC water

model specifies the point charge while the other two force fields do not. The second probable source of error is examined in the next section.

Ammonia/Acetone

While water is the most prevalent hydrogen bonding solvent, it is not the only one. Therefore, I have chosen to investigate how well the QM/MM method is able to treat ammonia as the hydrogen bonding solvent. The ammonia-acetone complex should be indicative of how well ammonia solvent interactions will be for carbonyl oxygen containing compounds. The hydrogen bonding conformer for the ammonia-acetone complex is shown in (Figure 3.16) while the binding energies and hydrogen bonding distances and angle are given in Table 3.17. The ammonia dimers will also provide additional insight in to the size of the error introduced in the water dimer from mixing the two force fields (CFF/SPC), since ammonia and the other molecules are both modeled using the CFF.



conformer 1

Figure 3.16: Ammonia/Acetone Complex

Table 3.17: Ammonia/Acetone (A/Ace) Dimer Results

	INDO	PM3	INDO/MM	MM/INDO
ΔE	20.4	3.42	3.43	4.25
R1	1.371	1.874	1.803	1.855
R2	2.457	2.879	2.807	2.898
A1	179.5	173.8	169.1	171.8

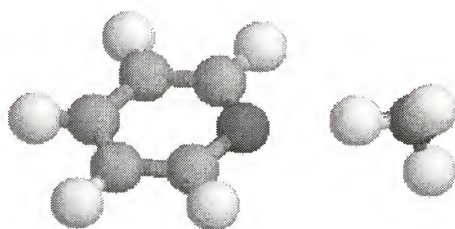
All Energies in kcal/mol
R1 = O(Ace) ... H(A) R2 = O(Ace) ... N(A)
A1 = O(Ace) ... H(A) - N(A)

The observed binding energy observed for the INDO/MM and MM/INDO models are more consistent than those obtained for the water - acetone complex. The 0.7 kcal/mol difference in binding energy is similar to the differences observed for the other test cases. The difference in the measured distances was ~ 0.05 Å while the difference in the angle was $\sim 3^\circ$. The increased consistency for the ammonia - acetone complex over the water - acetone complex indicates that it is not a deficiency of the force fields to model acetone, but rather a mismatch of the force fields is responsible for the inconsistency.

Ammonia/Pyridine

The ammonia/pyridine system is an important system to study for two reasons; first to test the interactions between an amino group and a conjugated nitrogen in a ring system, the second is to test the consistency of a single force field versus the two force field model used for describing the

water dimer test cases. The hydrogen bonding structure for this calculation is analogous to the structure used for the water/pyridine system (Figure 3.17). The predicted values for the binding energies and hydrogen bonding information are presented in Table 3.18.



conformer 1

Figure 3.17: Ammonia/Pyridine Complex

Table 3.18: Ammonia/Pyridine (A/P) Dimer Results

	INDO	PM3	INDO/MM	MM/INDO
ΔE	14.41	1.26	3.62	3.74
R1	1.476	3.042	1.920	1.957
R2	2.549	3.627	2.922	3.001
A1	172.7	118.4	169.6	172.5

All Energies in kcal/mol

R1 = N(P) ... H(A)

R2 = N(P) ... N(A)

A1 = N(P) ... H(A) - N(A)

The difference in the binding energy when comparing the INDO/MM and MM/INDO systems is ~ 0.1 kcal/mol and is similar to that observed for the water complex. The difference in the measured distances is ~ 0.05 Å opposed to 0.04 Å observed for the water complex. The results obtained from this section do not indicate that there is a substantial difference between using a single force field versus the two force field model.

Ammonia/Methanol

The ammonia-methanol system is also an important system for determining whether or not the ammonia solvent can be used for modeling the interactions between an amino group and an alcohol or organic acid. This system also serves as a test case for using an alcohol as the solvent as did the methanol - formamide complex. There are two hydrogen bonding conformations examined for this system (Figure 3.18). The calculated binding energies and hydrogen bonding information are given in Tables 3.19 and 3.20.



Figure 3.18: Ammonia/Methanol Complex

Table 3.19: Ammonia/Methanol (A/M) Dimer Results; Conformer 1

	INDO	PM3	INDO/MM	MM/INDO
ΔE	17.34	0.06	2.84	2.70
R1	1.362	1.876	1.846	1.838
R2	2.438	2.879	2.860	2.888
A1	178.3	173.5	177.9	178.8

All Energies in kcal/mol
 R1 = O(M) ... H(A) R2 = O(M) ... N(A)
 A1 = O(M) ... H(A) - N(A)

Table 3.20: Methanol/Ammonia (M/A) Dimer Results; Conformer 2

	INDO	PM3	INDO/MM	MM/INDO
ΔE	16.30	0.08	3.5	2.84
R1	1.476	1.928	1.950	1.998
R2	2.500	2.923	2.935	2.953
A1	178.5	168.2	169.2	175.1

All Energies in kcal/mol
 R1 = N(A) ... H(M) R2 = N(A) ... O(M)
 A1 = N(A) ... H(M) - O(M)

Comparing the binding distances from the INDO/MM and MM/INDO models shows that ~ 0.02 Å difference for conformer 1 and ~ 0.03 Å difference for conformer 2. These numbers are similar to those obtained for the water complex for conformer 1, although conformer 2 for the ammonia complex is more consistent. The same trend is observed when comparing the binding energies. The ammonia complex provides a more consistent representation for conformer 2 than the water complex. This is most likely an error introduced by the molecular mechanics force field.

Summary and Conclusions

The hybrid quantum mechanics-molecular mechanics (QM/MM) method developed herein has been shown to adequately model the hydrogen bonding interactions. This QM/MM method has demonstrated that it is capable of reproducing experimental results and results obtained from high level *ab initio* calculations. The QM/MM method developed herein has also been shown to be more consistent than previous QM/MM methods [120]. This increased consistency is largely due to correctly modeling the Coulombic interactions between the quantum mechanics region and the molecular mechanics region.

The larger error associated with the water - acetone complex and the second conformer of the water - methanol complex are believed to be an artifact of the molecular mechanics force fields. This error is due primarily to a mismatch of the two force fields used for modeling the molecular mechanics interactions. The two systems with the greatest differences between the INDO/MM and MM/INDO models had a negative binding energy for the MM/MM models. This error could be reduced by

choosing a more accurate force field to represent the molecular mechanic region, such as the MMx force fields by Allinger [67, 69].

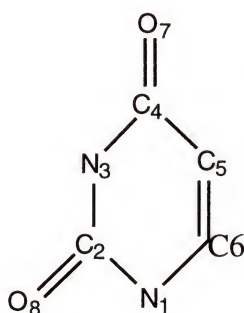
CHAPTER 4

APPLICATIONS OF THE QM/MM METHOD ON LARGE MOLECULAR SYSTEMS

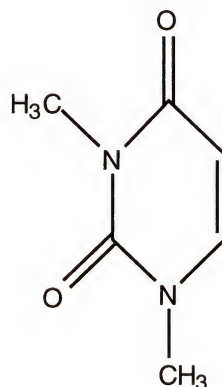
Uracil and Dimethyluracil Spectroscopy

Background

It has been shown by Becker and Kogan [194] that the luminescence characteristics of uracil (U) and 1,3-dimethyluracil (DMU) is different depending on the solvent. Both U and DMU fluoresce in polar hydrogen-bonding solvents (protic), but only DMU shows fluorescence in polar non-hydrogen bonding solvents (aprotic) and U only phosphors in the latter type of solvents.



uracil (U)



1,3-dimethyluracil (DMU)

Figure 4.1: Uracil and 1,3-Dimethyluracil Structures

Baraldi et al. [195] have conducted a theoretical study on the spectroscopy of U and several uracil derivatives using an INDO-based method called CS-INDO. The observed absorption spectra of uracil were well reproduced, and the photophysics were discussed. However, they did not include solvent interactions in the description and did not address the dependency of the emission on the solvent.

In a molecular dynamics (MD) study by Broo using classical potentials [196] the DMU complex in water was investigated. The MD simulation showed that each carbonyl oxygen of the DMU accepts protons via hydrogen bonding with at least two water molecules. The binding strength was not computed in that study, although the carbonyl oxygen with one neighboring CH₃ group (O7) bound a water molecule considerably more tightly than did the other carbonyl oxygen (O8).

Method

Molecular dynamics simulations have been performed on uracil and DMU with 98 water molecules using the hybrid INDO/MM method. The water molecules were included in the MM moiety, and the U, or DMU, was treated as the quantum system. Periodic boundary conditions were used, and the simulated system was enclosed in a 16x16x16 Å box. The system was first equilibrated at 77K by a combination of geometry optimization and MD simulations using a high-friction temperature scaling (λ) to remove all excess energy from the arbitrary initial structure. After equilibration, the system was run for 4.0 ps using a 0.1 fs time step, requiring approximately 6 hours of CPU time on an IBM RS6000/590 workstation.

Results

The hybrid MD simulation of the U-water system predicts a hydrogen-bonding pattern similar to that of the all classical MD simulation of the DMU-water systems. In Figures 4.2 and 4.3 the hydrogen bond distances between the two carbonyl oxygens and the neighboring water molecules are plotted as a function of the simulation time. The carbonyl oxygen with only one neighboring N-H group has two water molecules that are strongly hydrogen bonded during most of the simulation, (Figure 4.2). At the end of the simulation a third water molecule collided with one of the hydrogen bonded water molecules and exchanged position with the “permanent” water molecule close to the carbonyl oxygen. The second hydrogen binding site is less crowded during the simulation, (Figure 4.3). The difference in binding capability between the two sites was more pronounced in the pure classical MD simulation of DMU in water. Clearly at least one water molecule is stationary in the vicinity of the carbonyl oxygens of both U and DMU, at least during the time of our simulations. To investigate the solvent effects on the absorption spectrum of U and DMU several INDO/S-CIS calculations were performed. The absorption spectra were obtained from configuration interaction (CI), that involved all of the single excited configuration state functions that have $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ character (CIS). The absorption spectrum predicted by the INDO/S-CIS model, in vacuum, agrees well with the spectrum previously published by Baraldi [195] using a slightly different parameterization of the INDO model (CS-INDO). The effectiveness of the INDO/S model in reproducing the spectra of nucleotides is well documented [197-199].

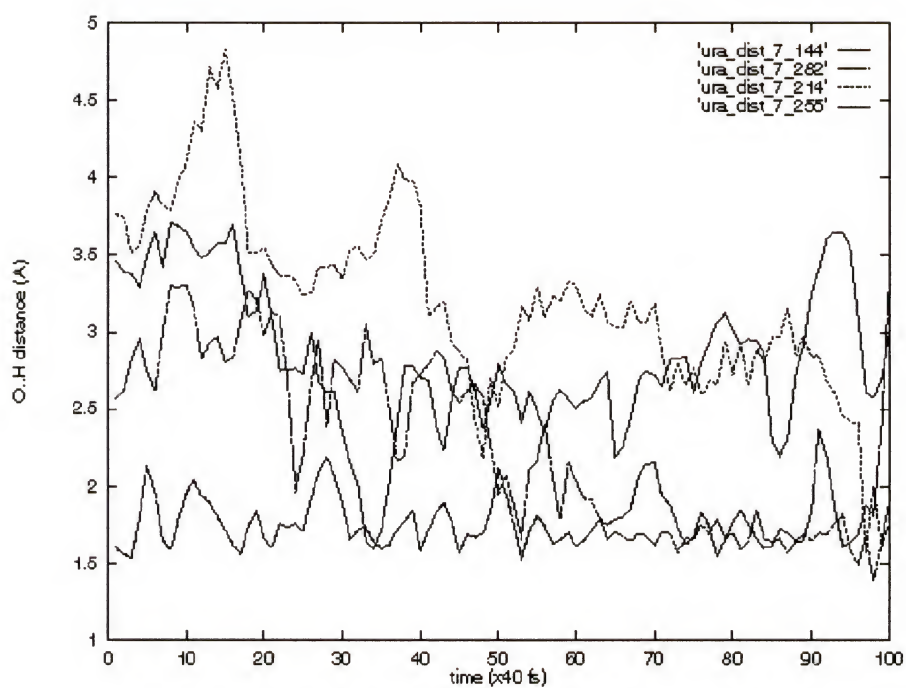


Figure 4.2: Hydrogen Bond Distance Between the O7 Carbonyl and the Neighboring Water Hydrogen's as a Function of Simulation Time.

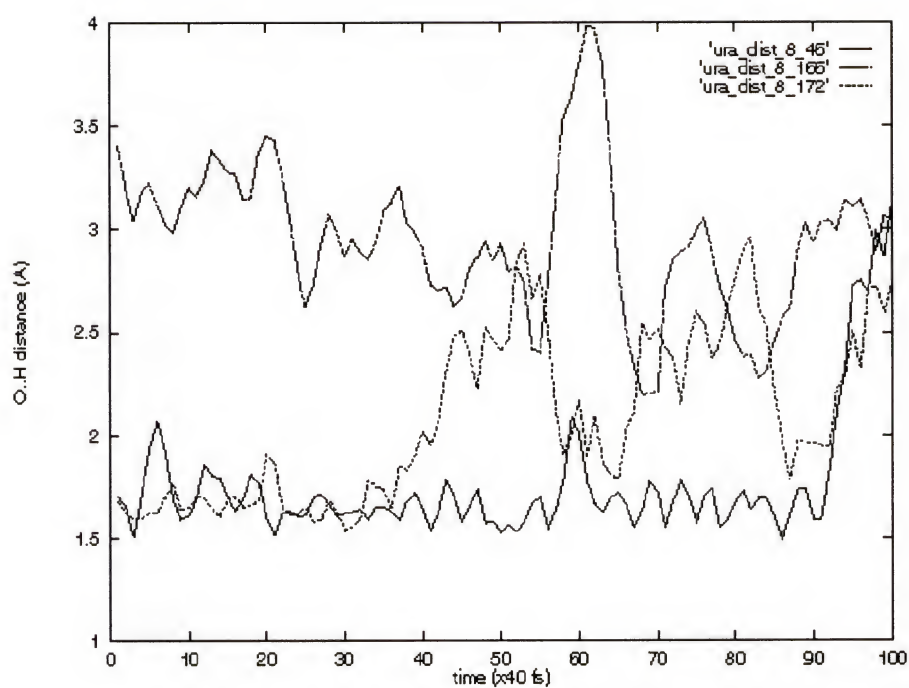


Figure 4.3: Hydrogen Bond Distance Between the O8 Carbonyl and the Neighboring Water Hydrogen's as a Function of Simulation Time.

Table 4.1: Comparison of Predicted Absorption Spectra of Uracil with Observed Spectra

	INDO	INDO +SCRF	INDO U + 8 H ₂ O	INDO U + 8 H ₂ O + SCRF	CS-INDO ref. 195	CASPT2 ref. 200	exp.
$n \rightarrow \pi^*$	32.3 0.0008	33.0 0.0008	33.2 0.0009	33.2 0.0009	32.7 0.001	36.6 0.0002	35.3^a
$n \rightarrow \pi^*$	38.4 0.000	38.4 0.0008	39.1 0.0004	38.7 0.0146	39.6 0.000	48.4 0.0000	41.7^{b,c}
$\pi \rightarrow \pi^*$	40.7 0.39	40.3 0.39	40.3 0.38	40.1 0.36	42.6 0.34	40.3 0.19	38.2^c
$\pi \rightarrow \pi^*$	48.8 0.20	47.4 0.21	48.0 0.21	48.1 0.20	50.4 0.19	46.9 0.08	46.5^c
$n \rightarrow \pi^*$	50.2 0.0002	50.6 0.0002	50.4 0.0008	50.3 0.0052	d	51.4 0.0000	d
$? \pi \rightarrow \sigma^*$	51.4 0.0003	52.2 0.0004	52.2 0.0005	52.0 0.0133	d	d	d
$\pi \rightarrow \pi^*$	52.5 0.08	53.5 0.11	53.0 0.10	53.2 0.11	52.2 0.47	52.1 0.29	51.3^c
$n \rightarrow \pi^*$	55.5 0.0007	54.7 0.0012	56.1 0.0035	56.3 0.0064	d	56.1 0.0001	d
$\pi \rightarrow \pi^*$	56.8 0.17	56.5 0.02	56.4 0.10	56.3 0.03	57.1 0.10	56.5 0.76	56.5^c
$^3(\pi \rightarrow \pi^*)$	15.9	16.6	16.1	16.0	24.4	d	22.2^e
$^3(\pi \rightarrow \pi^*)$	29.2	30.8	30.1	30.2	30.6	d	d

Energy is in 1,000 cm⁻¹; the nonbold numbers are oscillator strengths

a Electronic spectra of uracil in a supersonic jet [203].

b CD spectra of dioxuracil monophosphate in water solution [201]

c Uncertain assignment from the CD measurements

d This transition was not reported in the cited work

e Reported as phosphorescence in 2-MTHF solution at 77K [194]

Table 4.2: Comparison of Predicted Absorption Spectra of 1,3-Dimethyluracil with Observed Spectra.

	INDO	INDO +SCRF	INDO DMU + 8H ₂ O	INDO DMU + 8H ₂ O + SCRF	CS-INDO ref. 195	exp.
$n \rightarrow \pi^*$	30.7 0.0009	31.3 0.0009	30.8 0.0006	31.8 0.0009	31.7 0.000	35.3^a
$n \rightarrow \pi^*$	37.3 0.0001	37.3 0.0005	35.5 0.0013	37.2 0.0068	38.5 0.000	41.7?^{b,c}
$\pi \rightarrow \pi^*$	39.6 0.32	39.0 0.33	37.4 0.31	38.2 0.32	41.4 0.28	37.7^c
$\pi \rightarrow \pi^*$	47.8 0.20	46.7 0.21	44.3 0.23	46.0 0.22	48.1 0.16	46.5^c
$\pi \rightarrow \pi^*$	48.0 0.11	48.7 0.13	47.5 0.21	49.1 0.14	50.6 0.46	d
$n \rightarrow \pi^*$	49.4 0.0008	49.6 0.0003	49.6 0.0005	49.5 0.0006	d	d
$?\pi \rightarrow \sigma^*$	50.8 0.0002	51.5 0.0002		51.5 0.0045	d	d
$\pi \rightarrow \pi^*$	53.2 0.41	53.7 0.35	51.3 0.29	53.6 0.38	54.4 0.39	51.3^c
$n \rightarrow \pi^*$	54.6 0.0003	53.8 0.0002	53.4 0.0029	55.2 0.0080	d	d
$\pi \rightarrow \pi^*$	58.3 0.06	56.4 0.03	55.1 0.08	55.7 0.07	61.1 0.11	56.5^c
$^3(\pi \rightarrow \pi^*)$	17.2	17.6	16.2	16.4	24.4	d
$^3(\pi \rightarrow \pi^*)$	27.6	28.4	28.8	29.0	29.6	d

Energy is in 1,000 cm⁻¹; the nonbold numbers are oscillator strengths

a Electronic spectra of uracil in a supersonic jet [203].

b CD spectra of dioxyuracil monophosphate in water solution [201]

c Uncertain assignment from the CD measurements

d This transition was not reported in the cited work

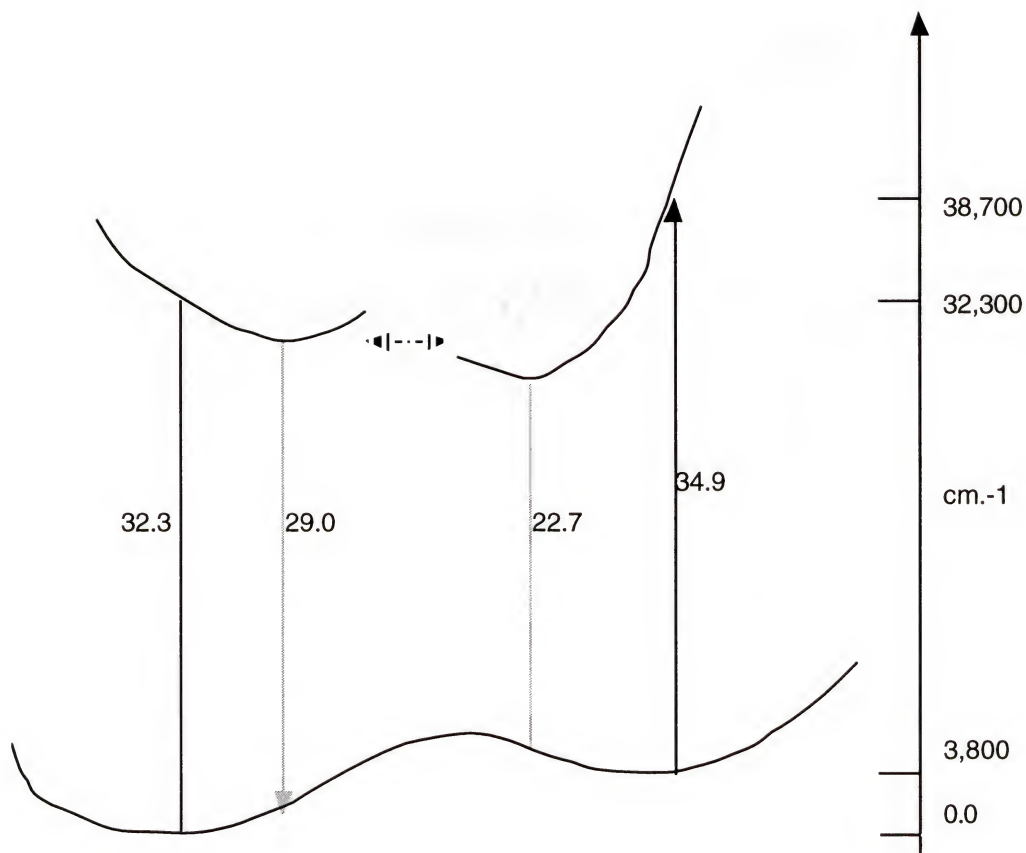
e Reported as phosphorescence in 2-MTHF solution at 77K [194]

The predicted absorption spectra of U and DMU in vacuum and in water solution are compared with the observed absorption spectra in Tables 4.1 and 4.2.

The solution spectra were computed using the traditional SCRF model using a spherical cavity [202]. Since the SCRF model does not include specific solvent interactions, we also computed the absorption spectra of U and DMU in the presence of eight water molecules. In the spectral calculation the water molecules were represented as point charges from the classical force field. Both U and DMU having four neighboring water molecules surrounding each of the carbonyl oxygens were geometry optimized using the QM/MM method utilizing the BFGS procedure. In both molecular complexes one water molecule was found to be hydrogen bonded to each carbonyl oxygen ($\text{O}\cdots\text{H}$) at a distance of approximately 1.65 Å. The remaining water molecules form a second solvation shell around the most tightly bound water molecule. In the $\text{U} + 8\text{H}_2\text{O}$ complex one water molecule forms a cyclic configuration with one hydrogen bond to O7 and a second hydrogen bond to the N3 hydrogen. This type of cyclic double-hydrogen-bonded complex cannot be formed with DMU and water. In this loosely bound system several energy minima are possible, and the above described geometry is just one of the many close lying conformers. The water structure around the carbonyl groups is similar for both U and DMU. The predicted spectra of these complexes are also found in Tables 4.1 and 4.2. All theoretical investigations predict the lowest state to be $n \rightarrow \pi^*$ in character, in agreement with the interpretation of the emission spectra reported by Becker et al. [194] and the supersonic jet spectrum of uracil reported by Fujii et al. [203].

Fujii observed a two band system, one with a transition located at 35,288 cm^{-1} (system I) and the other system with a transition located at 30,927 cm^{-1} (system II). The vibration progression of system I was found to be very similar to the ground state vibration spectrum, while the vibration progression for system II was very different. The 35,288 cm^{-1} peak was assigned as $n \rightarrow \pi^*$ and the 30,917 cm^{-1} peak was assigned to be an $n \rightarrow \pi^*$ transition from a different tautomer form of uracil. Fujii assigned the latter transition to come from the 2-enol tautomer based on an early *ab initio* calculation of the relative stability of the possible enol-keto tautomers of uracil, even though the suggested tautomer was calculated to be 72.1 kJ/mol less stable than the 2,4-diketo tautomer. Recently more sophisticated calculations of the relative stability of the different uracil tautomers have been reported [204,205]. The conclusion from the latter work is clear: only the 2,4-diketo tautomer of uracil is expected. The next stable tautomer is calculated to be less stable relative to the 2,4-diketo tautomer by 45 kJ/mol. A tempting alternative assignment of the two $n \rightarrow \pi^*$ peaks would be that both originate from the 2,4-diketo tautomer, and this reassignment is in reasonably good agreement with the INDO/S-CI and CS-INDO calculations. However, the CASPT2 predicted spectrum of uracil [200] would then be in error by about 0.6 eV for the $n \rightarrow \pi^*$ transitions. Upon inspection of the molecular orbitals (MO) from the INDO calculations two nonbonding MOs are identified. In both U and DMU the highest occupied nonbonding MO is slightly localized on the O7 carbonyl group, and the other occupied nonbonding MO is slightly localized at the other carbonyl group. Excitations from these two MO's will result in at least two low lying $n \rightarrow \pi^*$ transitions, which is predicted by all calculations. All calculations predict that the $n \rightarrow \pi^*$ state is the lowest emitting state [194,203,206].

In general, the emitting state is the lowest excited state of a given multiplicity, and this state is relaxed with respect to solvent and geometric parameters. Thus, our objective is to estimate the geometry of the lowest excited state in order to calculate the emission spectra of U and DMU. We have optimized the lowest excited state geometry of U by calculating the lowest open-shell singlet using the ROHF formalism [207,208]. Only small changes in geometry are found when comparing the $2^1A''$ and the $1^1A''$ states. The C4-O7 bond length is increased in the $2^1A''$ state, and there is a corresponding decrease of the C4-C5 bond length and increase of the C5-C6 bond length. The C4-O7 bond length suggests single-bond character in the $2^1A''$ state. Furthermore, the carbonyl oxygen bends toward the N3 hydrogen. These geometric changes suggest that a proton transfer might be favored in the excited state. If the proton transfer occurs, the emission could take place from two different tautomers. The proton transfer mechanism might offer an explanation of the two systems of bands observed in the supersonic jet fluorescence excitation spectrum reported by Fujii et al. [203]. The lowest $n \rightarrow \pi^*$ state of the 4-enol tautomer is located around $34,900\text{ cm}^{-1}$, which is slightly blue shifted with respect to the lowest $n \rightarrow \pi^*$ transition of the 2,4-diketo tautomer, (Figure 4.4). The lowest $\pi \rightarrow \pi^*$ transition of the 4-enol tautomer is calculated to be at $35,000\text{ cm}^{-1}$. Clearly, this spectrum does not agree well with observed spectrum recorded in the supersonic jet. The λ_{max} of the fluorescence of U, using the optimized excited state geometry, is predicted to be approximately $29,000\text{ cm}^{-1}$ in vacuum and approximately $30,200\text{ cm}^{-1}$ in a solution of water, modeled using the SCRF model.



Keto - tautomer <-----> enol tautomer

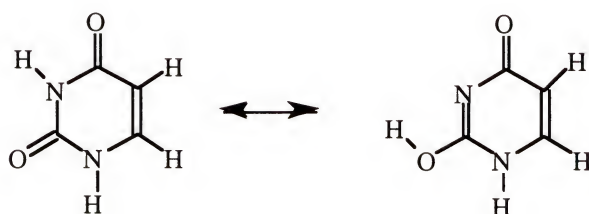


Figure 4.4: Schematic Drawing of the Keto-Enol Tautomerization

The energies of the excitations are from CIS calculation, and the keto-enol separation in the ground state of $3,800\text{ cm}^{-1}$ is from the *ab initio* calculations of ref. 204,205. The INDO/1 SCF calculations do not give barriers, but suggest that both keto and enol forms have real minima in both ground and excited states.

The observed λ_{max} for the fluorescence is 33,333 cm^{-1} in 2-MTHF and ethanol/methanol at 77 K [194] and 31,350 cm^{-1} in water solution [204,205] and the lowest triplet state of both U and DMU has $\pi \rightarrow \pi^*$ character. The $^3(\pi \rightarrow \pi^*)$ state is predicted to be much lower in energy when compared to what is observed for the phosphorescence of U in 2-MTHF solution. The $^3(\pi \rightarrow \pi^*)$ state is slightly blue shifted by the solvent, but the predicted transition energy is still far from the observed phosphorescence energy [209].

Let us now turn to the photophysics of the two molecules. It is commonly believed that if two states are close in energy, then vibrational coupling between the two states is facilitated and luminescence will be effected [210]. From Tables 4.1 and 4.2 it is clear that the second $n \rightarrow \pi^*$ state is predicted to be close in energy to the first $\pi \rightarrow \pi^*$ state. The energy separation of these two states is predicted to be 2,300 cm^{-1} for both U and DMU in vacuum. In water solution (SCRF) the state separation is decreased to 1,900 cm^{-1} for U and 1,700 cm^{-1} for DMU. With both reaction field and specific solvent interactions considered in the calculation, the energy splitting is 1,400 cm^{-1} for U and 1,000 cm^{-1} for DMU. Solvent interaction narrows the state energy difference considerably. Thus, we expect that vibrational motions of the skeleton will cause state mixing and offer an effective radiationless deactivation channel for both U and DMU in both types of solvents. To confirm the state-mixing mechanism for the deactivation, we have computed the absorption spectrum of DMU using 10 geometries obtained during the MD simulation of DMU and the surrounding water molecules. The predicted spectrum is depicted in Figure 4.5.

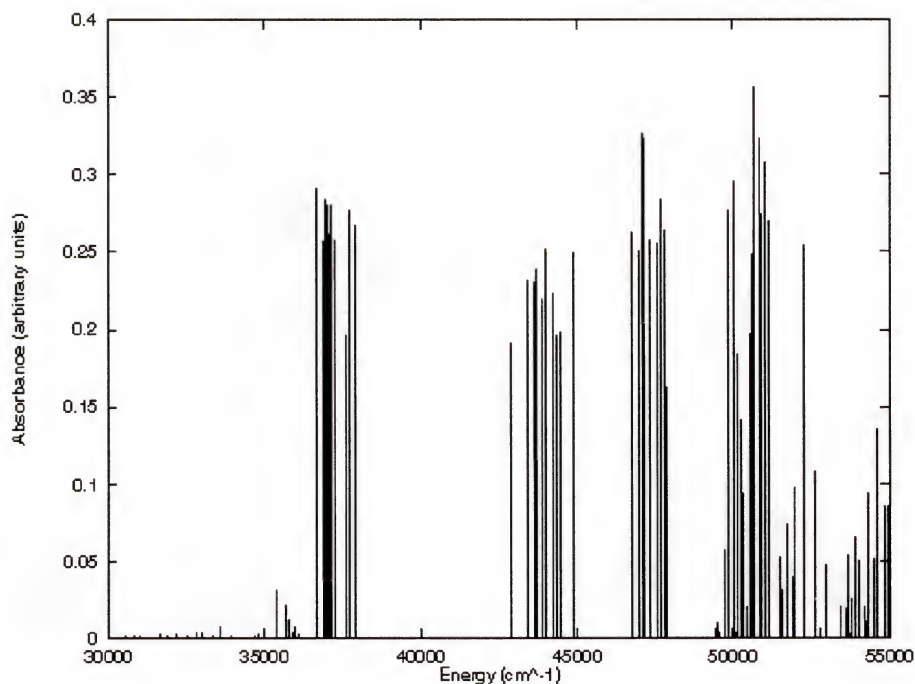


Figure 4.5: INDO/S-CIS Predicted Spectrum of 1,3-dimethyluracil

Spectrum is composite of 10 geometries taken from the 4ps MD simulation with a time separation of 0.4 ps.

Around $35,000\text{ cm}^{-1}$ a weak band appears in the predicted spectrum. By inspection of the CI eigenvectors the intensity increase of the $n \rightarrow \pi^*$ is due to mixing of the close lying $\pi \rightarrow \pi^*$ state. The vibrational mode responsible for the mixing is an out-of-plane bending mode of the O7 carbonyl group. Our results are in good accord with the observed very low quantum yield for the fluorescence. However, the reason for the absence of fluorescence and the large quantum yield for the phosphorescence for U in non-hydrogen-binding solvents is not explained by the above discussion. One possible explanation is that the N3 proton is transferred to the O7 carbonyl oxygen in

the excited state of U, (Figure 4.4). This proton transfer is blocked by a strong hydrogen bond between the solvent and U in hydrogen-bonding solvents. This type of proton transfer in the excited state is not possible for DMU. We predict the lowest triplet state of the 4-enol tautomer to be at $17,000\text{ cm}^{-1}$ and with $^3(\pi \rightarrow \pi^*)$ character, which is $\sim 5,000\text{ cm}^{-1}$ lower in energy than that observed experimentally. In the work by Becker and Kogan [195] the possible tautomerism was discussed, but their final conclusions regarding the difference in photophysical behavior for U and DMU most likely was resulted from a different microsolvent environment. The present study confirms that the solvation shell is different around U compared to DMU, but the difference is small and the effect on the calculated absorption spectrum is predicted to be small.

Summary

The hybrid QM/MM method has been used to investigate the dynamics and photophysics of uracil and 1,3-dimethyluracil in protic and aprotic solvents. The molecular dynamics simulation and geometry optimization of the uracil and 1,3-dimethyluracil in water solution demonstrated that at least one water molecule is tightly bound to each of the carbonyl oxygen atoms. The low quantum yield for the fluorescence is due to quenching by vibrational induced state mixing between one $n \rightarrow \pi^*$ and one $\pi \rightarrow \pi^*$ state that are predicted to be close in energy for both molecules. The active vibrational mode is an out-of-plane bending of the O7 carbonyl oxygen. In a non-hydrogen-bonding solvent uracil has a much larger quantum yield for phosphorescence than for fluorescence in any type of polar solvent. Since the dynamics for the two systems seems to be more or less the same, we suggest that the initial state for the phosphorescence is

the 4-enol tautomer of uracil. The 4-enol uracil is obtained by an internal proton transfer in the excited state that might be blocked by a hydrogen-bonded solvent molecule in the hydrogen-binding solvent.

Solvent Solute Charge Transfer

Introduction

It is well noted that the environment in which a molecule resides can greatly influence observed properties. It is therefore no surprise that proper modeling of solvent effects is crucial for predicting molecular electronic spectroscopy. The current methods for modeling solvent effects can be classified into two groups; implicit and explicit. The implicit treatment for solvent effects generally relies upon dielectric continuum models [211-214] and, the assumption that the “bulk” solvent effects are the same as the “local” solvent effects. Implicit solvent models have flourished recently, mainly due to their simplicity. Explicit solvent modeling methods treat each solvent molecule explicitly. This model has the major drawback of being computationally expensive, and usually requires using stochastic methods to properly reproduce the thermodynamics that are naturally included in the reaction field methods. Nevertheless such explicit consideration of the solvent is required whenever bonding between solute and solvent is greater than kT (thermal energy), and whenever there is charge transfer or exchange interactions.

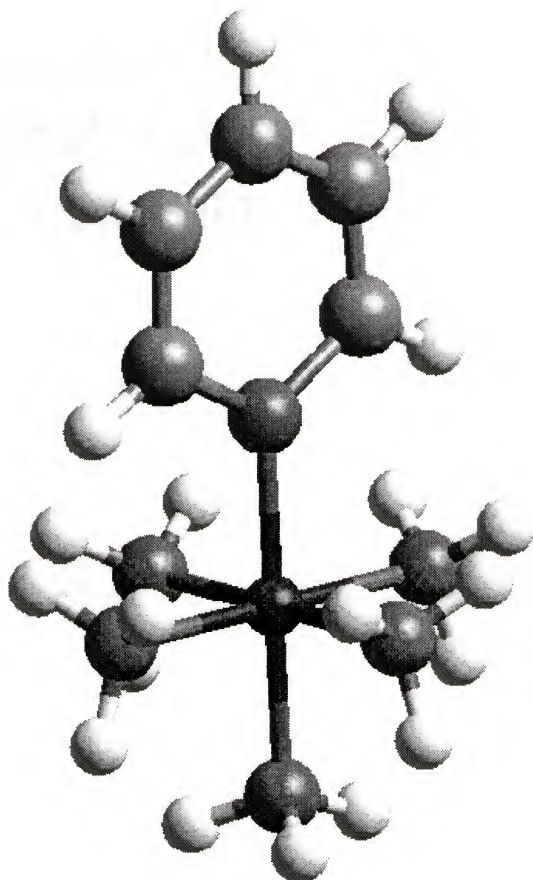


Figure 4.6: Ruthenium (II) Pentaaminopyridine Complex

Results

This work is a continuation of work done on the spectroscopy of ruthenium (II) pentaaminopyridine, $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$; (Figure 4.6), [215]. The original work was considered somewhat controversial. The initial study and conclusions drawn were based upon a limited number of

geometric configurations, and untested force fields. The N...O bond distance between the ammonia groups and their nearest neighboring water molecule was only 2.3 Å, compared to the experimental value of 2.8 Å. In this work, we examine the model described in this thesis to expose the elusive nature of the solvent sensitivity observed in the electronic absorption spectroscopy of $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ (py = pyridine) specifically [217-227] and, then confirm that the same physical phenomenon is responsible for the solvent sensitivity observed by the $[\text{Ru}(\text{NH}_3)_x\text{L}_{6-x}]^{2+}$ class of compounds [217-227].

The experimental low resolution solution (water) spectrum has two observable bands, the first around $25,000\text{ cm}^{-1}$ and the second around $41,000\text{ cm}^{-1}$ [216]. The first band appears to possess more intensity than the second band, although it is not a large difference. The INDO/1 gas phase calculation [65] agreed exceptionally well with the experiment transition energies, although the calculated oscillator strengths did not, (Table 4.3). The bands appearing around $25,000\text{ cm}^{-1}$ do not have any appreciable intensity, and the first band to have any intensity occurs at $37,000\text{ cm}^{-1}$. Additionally the relative intensities of the oscillator strengths were reversed when compared to the experimental spectrum. The gas phase calculation also predicted a $3,000\text{ cm}^{-1}$ splitting between the first two observed bands, compared to the $16,000\text{ cm}^{-1}$ observed experimentally [216].

A self-consistent reaction field was used to simulate the solvent ($\epsilon = 78.5$) [225]. The reaction field was able to get the proper ordering of the oscillator strength intensities, lowering the MLCT band, and was able to increase the splitting between the pyridine $\pi-\pi^*$ and MLCT to approximately $9,000\text{ cm}^{-1}$.

Table 4.3: Calculated Transition Energies for $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex.

Gas Phase			SCRF (water)			Experiment ^a		
E	Osc	Tran.	E	Osc	Tran	E	Osc	Tran
24.2	0.000	d-d	24.9	0.000	d-d	24.6	3.89	MLCT
24.9	0.000		25.0	0.000				
25.0	0.000		26.0	0.000				
33.1	0.000	d-d	32.9	0.000	d-d			
33.6	0.000		33.6	0.000				
33.8	0.000		35.0	0.002	MLCT			
37.5	0.070	$\pi-\pi^*$	35.3	0.230	MLCT			
40.5	0.200	MLCT	44.6	0.200	$\pi-\pi^*$	41.0	3.66	$\pi-\pi^*$

All energies in $1,000 \text{ cm}^{-1}$

a reference 216

This initial failure of the reaction field to predict the proper splitting, encouraged us to reexamine the model we used to describe the solvent. The next step was to take the $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ and 15 or 30 water molecules and treat the “supermolecule” using quantum mechanics. These geometry optimized “supermolecule” INDO/1 calculations were the basis of the first paper [215]. This model predicted a band splitting of $14,000 \text{ cm}^{-1}$ in reasonably good agreement with experiment, but the geometric configurations obtained using this model were questionable, as the $\text{N}\cdots\text{O}$ bond distances were approximately half an Ångström too small. Most important, however, was that the “supermolecule” results showed that

approximately 1.0 - 1.5 electrons were transferred from the solvent (waters) to the ruthenium complex. This led to the second controversial aspect of this initial work; was this a physical phenomenon or an artifact of INDO/1 approximations. In order to resolve some of the controversy created by the conclusions of paper I, we have continued to examine this problem.

To verify the charge transfer predicted by the INDO/1 calculation, an *ab initio* calculation [228] using the INDO/1 structure, for the ruthenium complex plus fifteen water molecules was performed. A 6-31G** basis set was used, and the calculated Mulliken populations [92] of the "supermolecule" is reported in Table 4.4.

Table 4.4: Comparison of the *ab initio* and INDO/1 Mulliken Charges for the $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex.

Model	net <i>ab initio</i>	net INDO/1	Ru ^a	Pyridine ^a	(NH ₃) ₅ ^a
Gas Phase	+2.000	+2.000	+0.375	+0.290	+1.335
Reaction Field	+2.000	+2.000	0.399	0.276	+1.325
Complex + 15 H ₂ O's	+0.853	+1.154	+0.383	+0.275	+0.496
Complex + 30 H ₂ O's	---	+1.103	+0.381	+0.260	+0.462
Complex + 10 Ethanes	---	+1.996	+0.471	+0.184	+1.341
Complex + 10 Acetylnitriles	---	+1.931	+0.620	+0.183	+1.128

a INDO/1 Mulliken Charges [92]

The charges obtained for the *ab initio* calculation showed that the INDO/1 charges were, if anything, too conservative and that the solvent-solute charge transfer is not an artifact of the INDO/1 approximations. It is also worth noting that it is the ammonia groups that are receiving the bulk of the charge transfer from the solvent. This is not surprising since the ruthenium cation has already experienced a charge transfer from the ammonia groups, even in the gas phase.

A hybrid quantum mechanics-molecular mechanics (QM/MM) calculation [229-233] was done to obtain several configurations of the ruthenium complex and the explicit water molecules. The QM/MM method was chosen for three reasons: first, modeling the ruthenium complex and over 200 water molecules explicitly using an *ab initio* theory is not computationally feasible at this time. Second, only the INDO/1 semi-empirical Hamiltonian is parameterized for transition metals and by itself, it models hydrogen bonds too strongly. The hybrid QM/MM method has been shown to model hydrogen bonds between the QM and MM regions accurately [233]. Finally, most molecular mechanics force fields model transition metals as having a formal charge of +2 or +3, while both the INDO/1 and *ab initio* calculations have shown that the formal charge on the ruthenium is less than + 0.5. These calculations have also shown that a sizable amount of polarization and charge transfer are also present in this system, and this cannot be modeled using present molecular mechanic force fields. A Monte Carlo method for sampling phase space was chosen. This simulation provided 500 uncorrelated structures, that were subsequently used for an INDO/1 CIS calculation. The CIS calculation used an "onion" model to represent the solvent. The first solvation shell was treated using quantum mechanics, while the remainder of the explicit

waters were treated as point charges. Finally, a self-consistent reaction field was placed around the entire supermolecule to remove boundary effects. The solvation shell was determined from the N...O radial distribution function for the five ammonia ligands, (Figure 4.7).

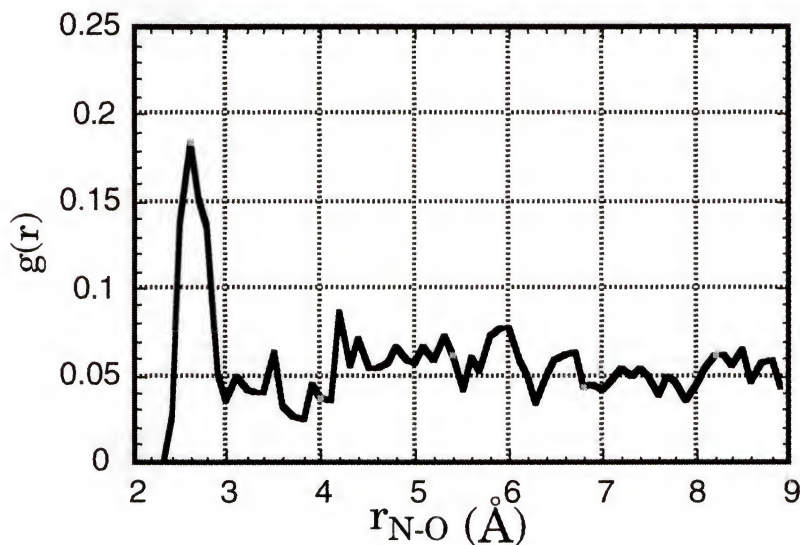


Figure 4.7: Radial Distribution Function for the Amino Groups and Surrounding Water Molecules.

The spectroscopic results from this simulation are given in Table 4.5, while the predicted spectrum may be seen in Figure 4.9. The calculated spectrum is in good agreement with the experimental spectrum for the first two observable transitions. The Mulliken charges from the QM/MM simulation also show that approximately one electron is transferred from the solvent to the solute.

Table 4.5: Calculated QMMM Transition Energies for the $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex.

Band	QM/MM		Experimental ^a	
	Energy (cm^{-1})	Osc	Energy (cm^{-1})	Osc
MLCT	24.2	0.994	24.6	3.89
$\pi-\pi^*$	37.0	0.611		
$\pi-\pi^*$	42.8	0.571	41.0	3.66

a reference 216

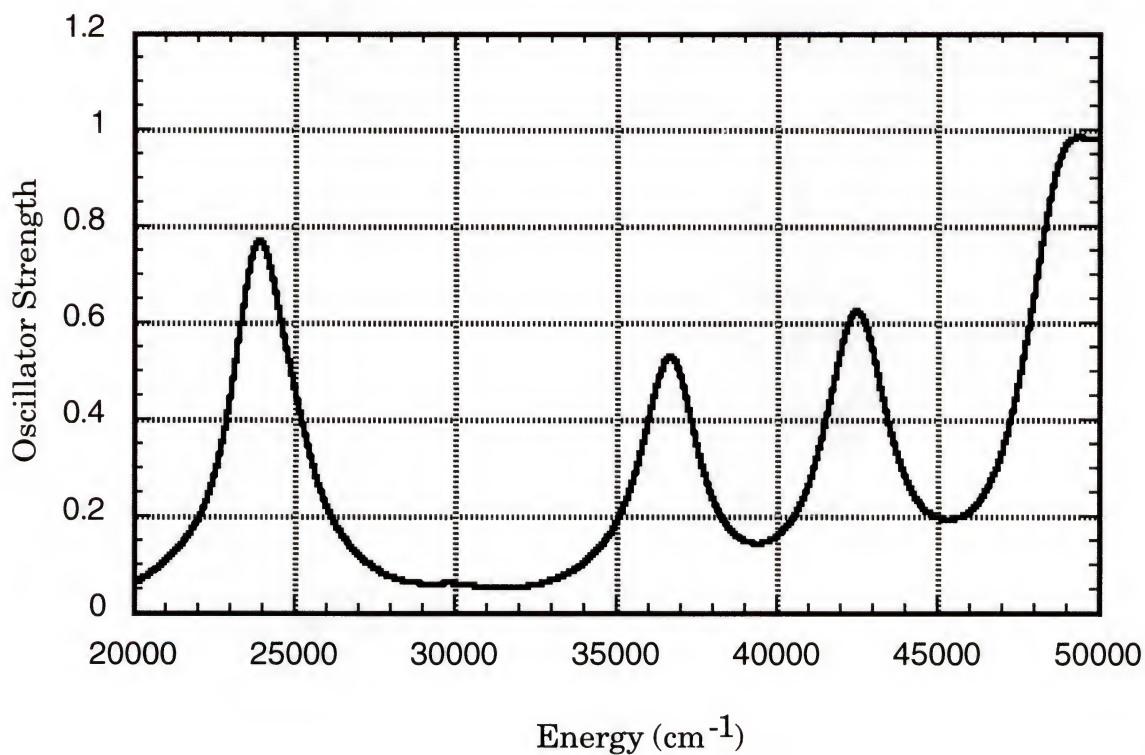


Figure 4.8: QM/MM Predicted Electronic Spectra for $[\text{Ru}(\text{NH}_3)_5\text{py}]^{2+}$ Complex.

Two points of interest when comparing the experimental and QM/MM calculated spectra: first, the experimental spectra was not measured past $41,000\text{ cm}^{-1}$, although it does indicate that a third peak is starting in that region. Second, the QM/MM calculated spectra for the +3 isomer was also in agreement with experiment, in that the first observable peak occurs around $37,000\text{ cm}^{-1}$.

Summary and Conclusions

In conclusion, we have shown that the large MLCT splitting arises from the charge transfer from the solvent to the solute. This is shown by comparison of the Mulliken charges for the ruthenium complex in the different explicit solvents. This is most likely the same source of the solvent sensitivity for the $[\text{Ru}(\text{NH}_3)_x\text{L}_{6-x}]^{2+}$ class of compounds. It is also our believe that most molecular mechanics force fields should be reevaluated to model a more realistic charge for metals, since a +2 or +3 formal charge is rare in quantum chemical calculations where the charge on the metals is between +0.5 and +1.0.

This work also shows the hybrid quantum mechanics-molecular mechanics method is capable of modeling systems containing metals, while standard molecular mechanics packages are not parameterized for most transition metals. The QM/MM method is also capable of providing insight into the electronic structure for systems that cannot be achieved using molecular mechanics.

Dissociation Energies

Introduction

The ability to make and break covalent bonds is one of the many advantages that quantum mechanics has over molecular mechanics models. Unfortunately, quantum mechanics calculations are slow when compared to molecular mechanics and therefore cannot be used to model large systems, such as enzymes, proteins, and molecules in solution. The hybrid quantum mechanics-molecular mechanics (QM/MM) models allows one to overcome the “pitfalls” of molecular mechanics while maintaining a relatively fast computational method. The QM/MM model maintains the ability of quantum mechanics to make or break covalent bonds, provided that these bonds are contained within the quantum mechanics partition, while utilizing molecular mechanics methods to model the solvent or protein environment. This focus of this work is to investigate how well the QM/MM method can model the dissociation energy for hydrogen halides (HX) in aqueous solution.

The potential energy surface for HF in aqueous solution is expected to have a bound minima that is lower in energy than that obtained for the dissociated species, since HF is a weak acid and does not ionize 100% in water [9, 14]. The potential energy surfaces for HCl and HBr in aqueous solution is expected to either possess a shallow bonding minimum or favor the dissociated species in solution, since these are strong acids and ionize 100% in aqueous solution. The dissociation of these acids provides an ideal

model for testing the QM/MM methods ability to model bond dissociation and formation.

Results

The results obtained using the QM/MM method are compared to the results obtained from gas phase calculations and the results obtained using both a spherical and elliptical cavity self-consistent reaction field (SCRF) calculation. The results for the gas phase and reaction field calculations were obtained using single point energy calculations at the specified H...X distances. The results from the QM/MM method obtained from a Monte Carlo simulations. These simulations held the H...X bond distances constant, while allowing the water molecules to move freely. The binding energies for the QM/MM system are calculated by using the average total energy from the simulation and subtracting the energy of the H...X molecule, geometry optimized structure, and then subtracting off the average total energy from a Monte Carlo simulation of 251 water molecules, the same number of water molecules used in the QM/MM simulation. The numbers for the HF, HCl, and HBr calculations are reported in Tables 4.6 - 4.8, while the potential energy curves are given in Figures 4.9 - 4.11. As you can see the QM/MM results for the bond energies from 1.0 Å to 2.0 Å is not noticeably different from those obtained using the SCRF elliptical cavity calculations. The QM/MM results start diverging from the SCRF results at H...X distances greater than 2.0 Å. This divergence of the two methods is believed to be mostly due to the inability of the self-consistent reaction field (SCRF) to adequately model the solvent accessible surface using a simple sphere of ellipse.

Table 4.6: HF Binding Energy

HF Distance (Å)	Gas Phase (kcal/mol)	SCRF 1 ^A (kcal/mol)	SCRF 2 ^B (kcal/mol)	QM/MM ^C (kcal/mol)
0.8	-5,715.3	-5,722.8	-5,732.9	-5,730.6
0.9	-5,784.3	-5,792.2	-5,802.1	-5,798.7
1.0	-5,803.8	-5,812.0	-5,821.9	-5,820.6
1.1	-5,794.8	-5,803.6	-5,813.6	-5,808.7
1.2	-5,770.7	-5,780.2	-5,790.5	-5,788.7
1.3	-5,742.4	-5,750.0	-5,760.9	-5,758.2
1.4	-5,720.6	-5,724.1	-5,729.9	-5,729.7
1.5	-5,704.6	-5,706.1	-5,707.4	-5,712.9
1.6	-5,693.0	-5,693.7	-5,694.1	-5,700.9
1.7	-5,684.9	5,685.2	-5,685.3	-5,694.1
1.8	-5,679.2	-5,679.3	-5,636.9	-5,688.5
1.9	-5,675.2	-5,675.3	-5,625.1	-5,682.1
2.0	-5,672.5	-5,601.2	-5,617.7	-5,676.3
2.1	-5,670.7	-5,602.8	-5,613.9	-5,673.4
2.2	-5,669.4	-5,610.4	-5,612.7	-5,675.2
2.3	-5,668.6	-5,622.1	-5,613.4	-5,674.5
2.4	-5,668.0	-5,636.9	-5,615.3	-5,670.2
2.5	-5,667.7	-5,653.9	-5,618.0	-5,670.8
3.0	-5,667.0	xxx	xxx	-5,669.5
3.5	-5,666.9	xxx	xxx	-5,674.1
4.0	-5,666.9	xxx	xxx	-5,680.3
4.5	-5,666.9	xxx	xxx	-5,679.8
5.0	-5,666.9	xxx	xxx	-5,681.6

A C1 theory using spherical cavity

B C1 theory using elliptical cavity

C QM partition only H-F, all waters classical

Table 4.7: HCl Binding Energy

HCl Distance (Å)	Gas Phase (kcal/mol)	SCRF 1 ^A (kcal/mol)	SCRF 2 ^B (kcal/mol)	QM/MM ^C (kcal/mol)
1.0	-1,744.3	-1,751.1	-1,772.8	-1,760.9
1.1	-1,817.3	-1,824.6	-1,846.9	-1,833.9
1.2	-1,856.4	-1,864.1	-1,887.5	-1,875.3
1.3	-1,872.5	-1,880.7	-1,905.6	-1,893.1
1.4	-1,873.4	-1,882.3	-1,909.2	-1,895.1
1.5	-1,864.8	-1,874.5	-1,904.1	-1,887.2
1.6	-1,851.1	-1,861.2	-1,894.2	-1,876.4
1.7	-1,838.8	-1,845.6	-1,882.5	-1,860.7
1.8	-1,828.6	-1,832.4	-1,870.9	-1,865.6
1.9	-1,820.6	-1,822.6	-1,860.7	-1,829.3
2.0	-1,814.4	-1,815.4	-1,852.4	-1,820.1
2.1	1,809.8	-1,810.2	-1,846.4	-1,815.0
2.2	-1,806.3	-1,806.5	-1,842.4	-1,811.7
2.3	-1,803.7	-1,803.8	-1,840.2	-1,808.7
2.4	-1,801.9	-1,801.9	-1,839.4	-1,808.1
2.5	-1,800.6	-1,758.0	-1,839.8	-1,802.8
3.0	-1,798.0	xxx	xxx	
3.5	-1,797.6	xxx	xxx	-1,800.8
4.0	-1,797.5	xxx	xxx	-1,804.5
4.5	-1,797.5	xxx	xxx	-1,802.2
5.0	-1,797.5	xxx	xxx	-1,805.7

A C1 theory using spherical cavity

B C1 theory using elliptical cavity

C QM partition only H-Cl, all waters classical

Table 4.8: HBr Binding Energy

HBr Distance (Å)	Gas Phase (kcal/mol)	SCRF 1 ^A (kcal/mol)	SCRF 2 ^B (kcal/mol)	QM/MM ^C (kcal/mol)
1.0	-1,638.3	-1,641.7	-1,658.7	-1,654.9
1.1	-1,742.6	-1,746.0	-1,763.0	-1,761.5
1.2	-1,807.9	-1,811.4	-1,828.7	-1,825.8
1.3	-1,845.4	-1,849.0	-1,866.7	-1,863.8
1.4	-1,863.1	-1,866.9	-1,885.4	-1,882.9
1.5	-1,867.2	-1,871.2	-1,891.0	-1,887.8
1.6	-1,862.3	-1,866.6	-1,888.0	-1,885.4
1.7	-1,851.7	-1,856.3	-1,880.0	-1,876.2
1.8	-1,840.8	-1,843.9	-1,869.5	-1,863.3
1.9	-1,831.3	-1,833.3	-1,858.2	-1,849.7
2.0	-1,823.6	-1,824.7	-1,847.4	-1,836.5
2.1	-1,817.3	-1,817.9	-1,837.9	-1,825.1
2.2	-1,812.4	-1,812.8	-1,830.1	-1,815.8
2.3	-1,808.7	-1,808.9	-1,824.1	-1,814.5
2.4	-1,805.9	-1,806.0	-1,819.9	-1,813.4
2.5	-1,803.8	-1,803.9	-1,817.2	-1,810.9
3.0	-1,799.5	xxx	xxx	-1,802.9
3.5	-1,798.8	xxx	xxx	-1,803.6
4.0	-1,798.7	xxx	xxx	-1,809.4
4.5	-1,798.7	xxx	xxx	-1,815.2
5.0	-1,798.7	xxx	xxx	-1,812.4

A C1 theory using spherical cavity

B C1 theory using elliptical cavity

C QM partition only H-Br, all waters classical

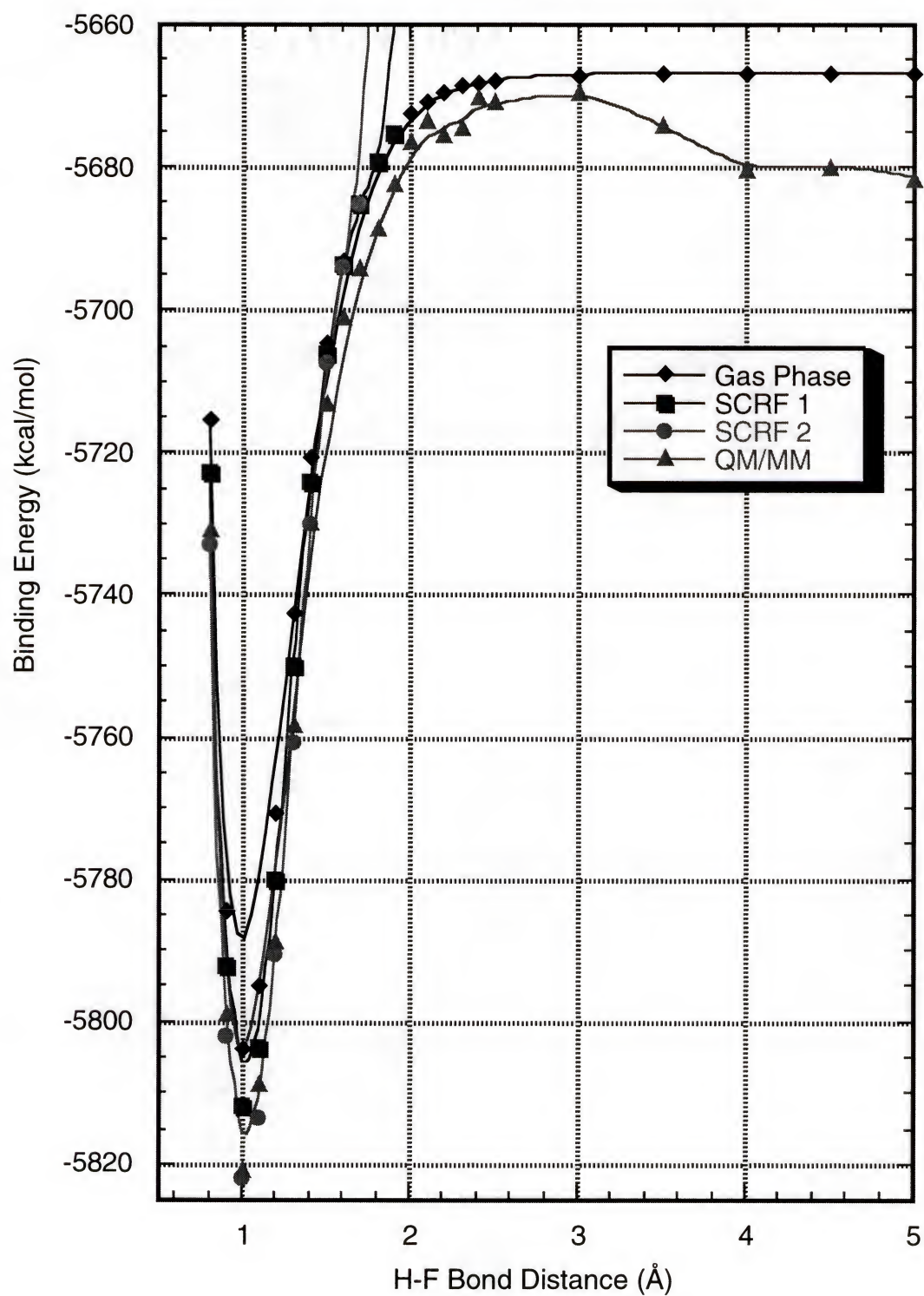


Figure 4.9: HF Potential Energy vs/ Bond Distance

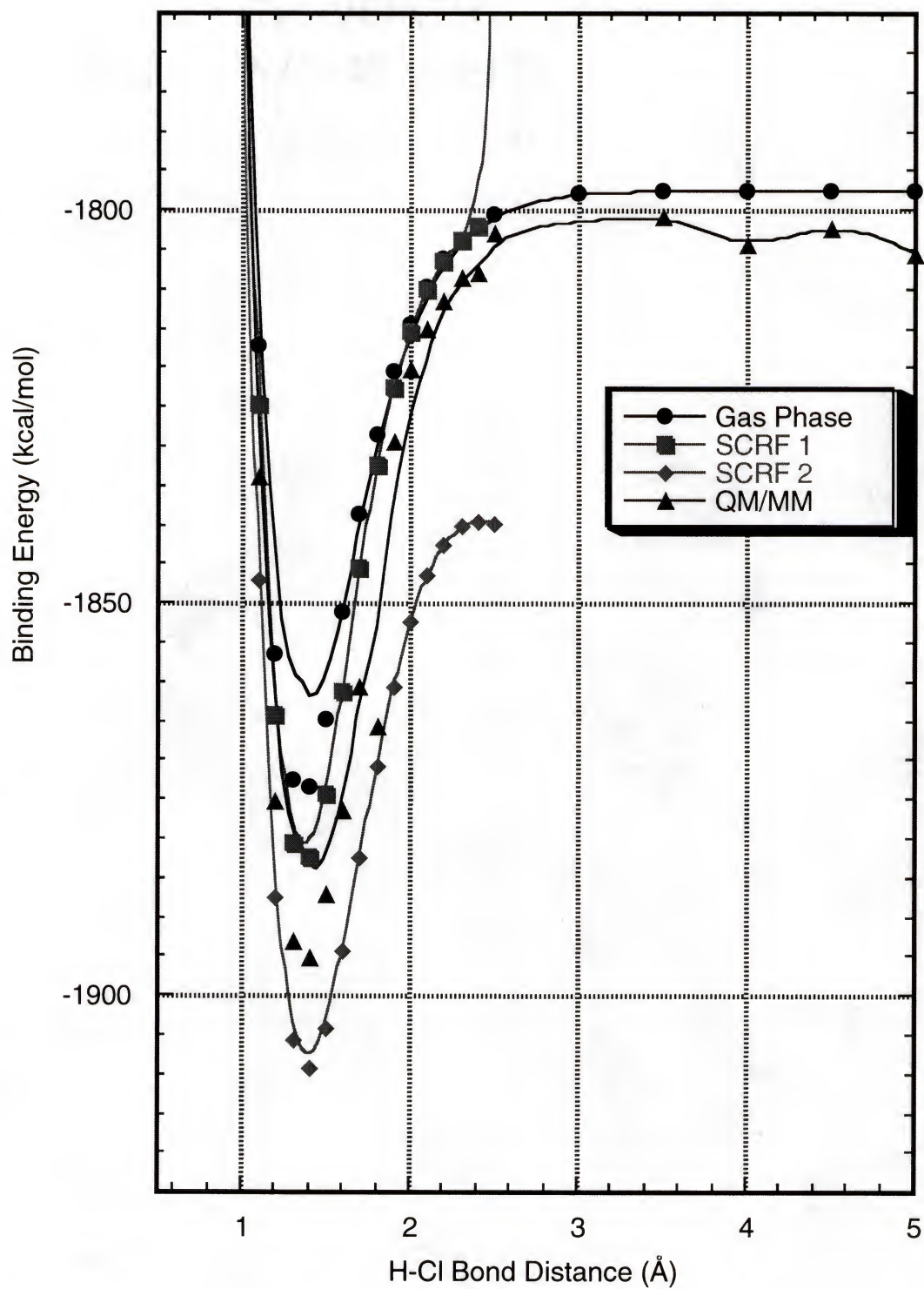


Figure 4.10: HCl Potential Energy vs/ Bond Distance

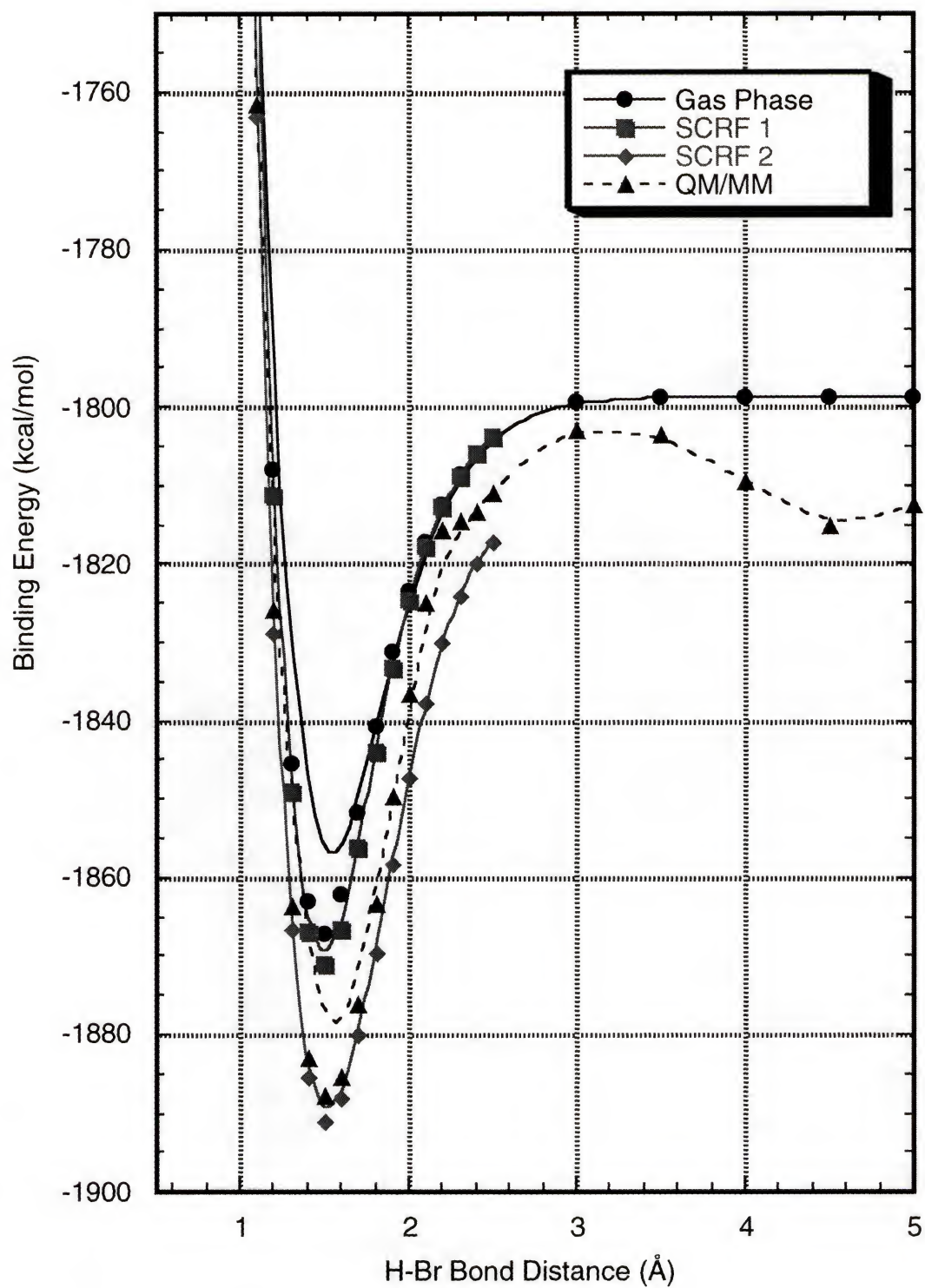


Figure 4.11: HBr Potential Energy vs/ Bond Distance

For the HF example the barrier between the bound state and the unbound state is ~210 kcal/mol for the SCRF methods while the QM/MM procedure has reduced this to ~150 kcal/mol. For the other examples we see that the binding energy for HCl is ~ 95 kcal/mol and is ~85 kcal/mol for HBr using the QM/MM method. These numbers correspond quite well with the experimental bond energies, (Table 4.9).

Table 4.9: H...X Bond Energies

	Exp ^A	QM/MM ^B	QM/MM ^C	QM/MM ^D
HF	-135.0	-151.1	-143.7	-142.2
HCl	-103.0	-94.3	-116.2	-108.2
HBr	-87.0	-84.9	-52.4	-83.1

A General Chemistry; Petrucci and Harwood, 7 ed. 1997

B INDO/SPC

C AM1/SPC

D PM3/SPC

The INDO and PM3 Hamiltonians predict the bond energy better than the AM1 Hamiltonian. This could be in part since the QM/MM method used was developed for the INDO Hamiltonian and has not been readjusted to compensate for AM1 or PM3 Hamiltonians. This does not appear to be a problem with the geometries obtained from these method since the F...H and H...O radial distribution functions for all three Hamiltonians are within 0.1 Å agreement for the distance of the first solvation shell.

Although these results are in good agreement with experiment the large standard deviation observed at the longer H...X bond distances and

the oscillating energies led to further investigation of the long range calculations. The source of the odd behavior at long dissociation energies is believed to originate from the interaction of solvent (classical mechanics) molecules coming between the two nuclei of the solute (quantum mechanics) molecule, (Figure 4.12). This type of interaction is not limited to solvent molecules, this could also occur when modeling large proteins or receptor binding sites when they are modeled using a QM/MM method.

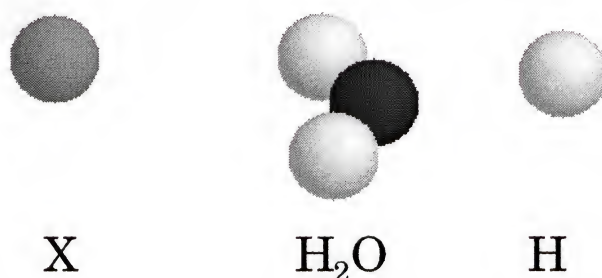


Figure 4.12: Solvent Interactions at Long H...X Dissociation Distances

Although the QM/MM method does not crash when these types of interactions occur, the design of the QM/MM method does not properly account for the charge screening and charge transfer that occurs from the solute (HX) and the solvent (H₂O) when the solvent comes between two solute nuclei. In order to test the significance of these types of interactions single point energy calculations have been performed on while varying the H...X bond distances while placing one of the H-O bonds from a water molecules between the H...X. Comparing the binding energies calculated using the QM/MM model to the full QM model will enable us to determine how significant these interactions are. The structures used for this test are

similar to conformation seen in Figure 4.12, while the binding energies for the QM/MM and QM models are provided in Tables 4.10 - 4.12.

Table 4.10: Difference in HF Binding Energy When Comparing QM and QM/MM Models

HF Distance (Å)	QM (kcal/mol)	QM/MM (kcal/mol)	Error (kcal/mol)
4.0	-5,707.0	-4,489.6	1,217.4
4.5	-5,686.4	-4,512.1	1,174.3
5.0	-5,675.3	-4,515.1	1,160.2
5.5	-5,670.4	-4,515.3	1,155.1
6.0	-5,668.2	-4,147.8	1,520.4
6.5	-5,667.4	-4,139.1	1,528.3
7.0	-5,106.0	-3,419.7	1,686.3
10.0	-5,253.9	-4,107.6	1,146.3
Ave Error			1,323.5

Table 4.11: Difference in HCl Binding Energy When Comparing QM and QM/MM Models

HCl Distance (Å)	QM (kcal/mol)	QM/MM (kcal/mol)	Error (kcal/mol)
4.0	-1,827.9	-617.0	1,210.9
4.5	-1,814.2	-642.0	1,172.2
5.0	-1,806.0	-645.5	1,160.5
5.5	-1,801.4	-645.5	1,155.9
6.0	-1,799.1	-645.5	1,153.6
6.5	-1,798.1	-645.5	1,152.6
7.0	-1,797.6	-645.5	1,152.1
10.0	-1,520.5	-373.9	1,146.6
Ave Error			1,163.1

Table 4.12: Difference in HBr Binding Energy When Comparing QM and QM/MM Models

HBr Distance (Å)	QM (kcal/mol)	QM/MM (kcal/mol)	Error (kcal/mol)
4.0	-1,818.2	-614.5	1,203.7
4.5	-1,812.8	-642.2	1,170.6
5.0	-1,806.8	-646.4	1,160.4
5.5	-1,802.7	-647.0	1,155.7
6.0	-1,800.5	-647.0	1,153.5
6.5	-1,799.4	-647.0	1,152.4
7.0	-1,798.9	-647.0	1,151.9
10.0	-1,798.5	-214.7	1,583.8
Ave Error			1,216.5

The ~1200 kcal/mol difference in the binding energy obtained from the QM and QM/MM simulations is somewhat concerning, especially when predicting binding energies or dissociation energies. This limitation would significantly reduce the usefulness of the QM/MM methods, since the first solvent shell would have to be included in the QM region. Alternatively, the QM region is restricted to systems that do not have any cavities that a solvent molecule could enter. Another method for calculating the binding energies with the QM/MM method is to use the configurations obtained from the dynamics or Monte Carlo simulation and perform a quantum mechanics calculation on the solute and nearest neighboring water molecules for the uncorrelated structures. This method is similar to the method used for obtaining the theoretical spectra in the previous section. The results for the improved QM/MM region are presented in Table 4.13, while the potential energies are plotted in Figure 4.13.

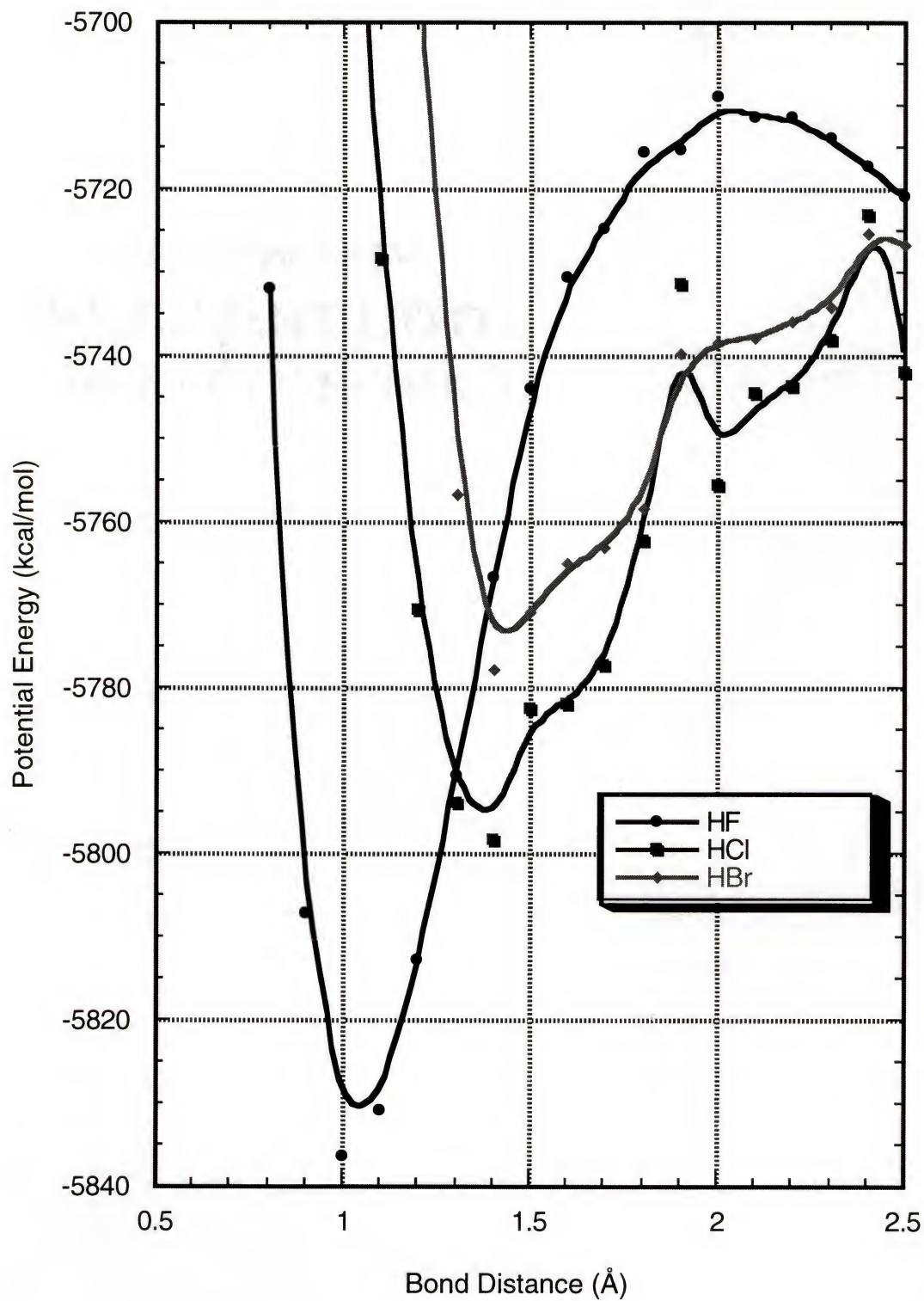


Figure 4.13: Dissociation Energy of HX Dimers + 4 QM H₂O's

Table 4.13: QM/MM Dissociation Energy of HX Dimers with 4 QM H₂O's

H...X Distance (Å)	HF (kcal/mol)	HCl (kcal/mol)	HBr (kcal/mol)
0.8	-5,731.8	xxx	xxx
0.9	-5,807.1	xxx	xxx
1.0	-5,836.4	-1,747.7	-1,629.4
1.1	-5,830.9	-1,828.2	-1,738.1
1.2	-5,812.8	-1,870.5	-1,795.8
1.3	-5,790.4	-1,893.8	-1,856.5
1.4	-5,766.6	-1,898.3	-1,877.7
1.5	-5,743.9	-1,882.5	-1,870.7
1.6	-5,730.4	-1,881.9	-1,865.1
1.7	-5,724.7	-1,877.3	-1,863.0
1.8	-5,715.6	-1,862.3	-1,858.3
1.9	-5,715.2	-1,831.3	-1,839.7
2.0	-5,708.9	-1,855.6	-1,838.2
2.1	-5,711.3	-1,844.5	-1,837.7
2.2	-5,711.4	-1,843.7	-1,835.8
2.3	-5,713.8	-1,838.1	-1,834.2
2.4	-5,717.3	-1,822.9	-1,825.3
2.5	-5,720.9	-1,841.9	-1,826.5

The results obtained from the larger QM region energy calculations have reduced the bond dissociation energies for all three compounds. This is in part due to inclusion of the solvent polarization and charge screening. The potential energies for HCl and HBr show some oscillation in the binding energies. This oscillation is associated with the binding energies between the water molecules. The size of the QM region was limited to four water molecules. The approximation worked relatively well for HF, although the larger halogen nuclei reduced the binding energies between the water molecules on average. These results could be improved further by increasing the QM region to include 10 - 20 water molecules, although this would increase the CPU time required quite substantially.

Summary

The QM/MM method has some severe limitations that must be addressed, especially when the QM region is not self contained. The QM/MM model is capable of searching phase space, although the potential energies are not that reliable unless care is taken to properly choose the QM region. These small acids should be one of the more difficult problems for the QM/MM method, since they disassociate into small ions. This could be a problem since the QM/MM interface is modeled using electrostatic interactions only. The errors obtained from using a small QM region appear to be somewhat constant. This indicates that these interactions might be parameterizable.

The errors observed in the QM/MM procedure appear to be independent of the QM Hamiltonian chosen. The AM1 Hamiltonian, for this QM/MM implementation, does not model the QM/MM energetic interactions properly, although the geometries obtained are similar to both

the INDO and PM3 results. The PM3 results are consistently lower in energy than experimental results, the error associated using this QM method could be improved by adjusting the Lennard-Jones (12-6) parameters, as has been done by most QM/MM methods [120].

CHAPTER 5 EMPIRICAL BAND WIDTH STUDY

Introduction

Absorption spectroscopy has become a widely utilized tool for chemical analysis and characterization of molecules. While there are several different types of absorption spectroscopy we are primarily interested here in UV/Visible spectroscopy, which is generated through electronic excitations.

Observed spectroscopic peaks are characterized by their transition frequency (ν), intensities or oscillator strength (f_{osc}), and band shape.

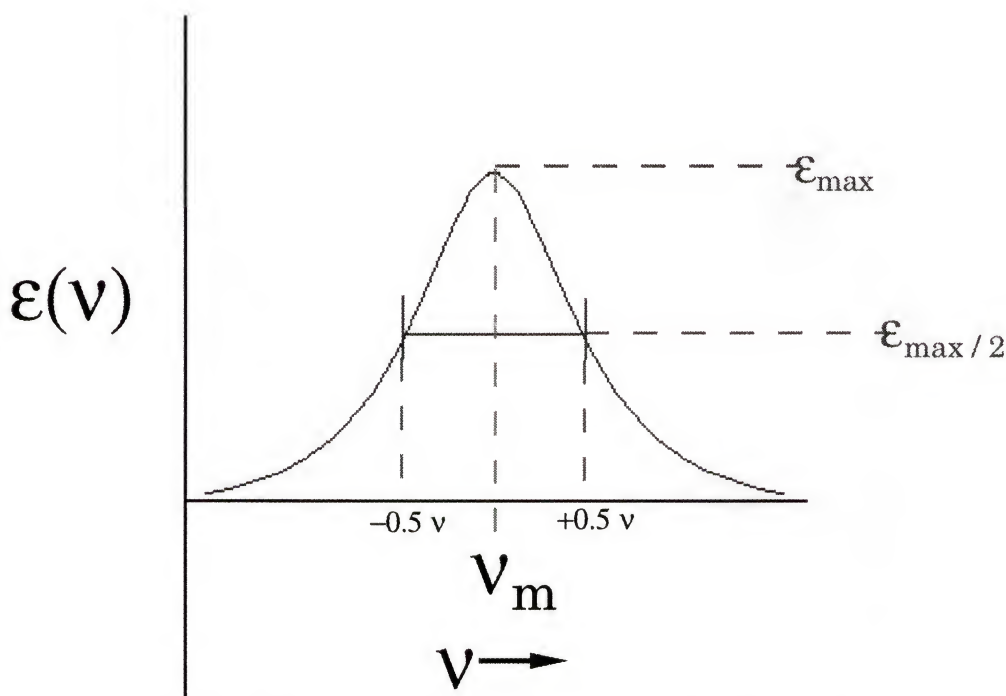


Figure 5.1: Anatomy of an Ideal Spectral Peak

Although each of these properties can provide valuable information for characterization of a molecule, it has been generally accepted that the frequency is the most important property for chemical analysis. For this reason theory has been aimed principally towards calculating the transition energy, and several successful methods have been developed [234-246]. While transition strength and shape have not been considered as vital for chemical analysis, they have certainly not been completely neglected [40, 234-236, 247-262]. Very early it was recognized that the intensity of a transition can be as important in assigning spectra as transition energy, and this is especially so when close lying bands are calculated in the reverse order, but transition strength and its polarization can provide definitive proof. The methods available for predicting band shapes are often restricted to small systems as they generally involve statistical mechanics [40, 254-262], or use an accurate treatment of the potential and attempt to solve the nuclear problem on this potential [247-251].

The focus of this work is to predict the band shape for molecular electronic absorption spectroscopy for large systems in condensed phases [252]. Under such conditions there are many phenomena that broaden transitions. Perhaps the major contributors for the broadening is the unresolved transitions to the various vibrational levels of the excited state. Vibrational broadening is not the only culprit though, because there are also excitations to the various rotational levels and anisotropic interactions with the medium (solvent) that are also responsible for line broadening. The excitations to the various rotational levels are also dependent upon

temperature, while the anisotropic interactions with the solvent are primarily dependent upon the solute and solvent. Besides line broadening there are also several factors that will affect the appearance of the band shape [252]. These factors include overlapping electronic transitions along with various perturbations between the excited states, such as vibronic and spin-orbit coupling. The effects to the appearance of the spectrum that arises from overlapping transitions are plotted in Figure 5.2.

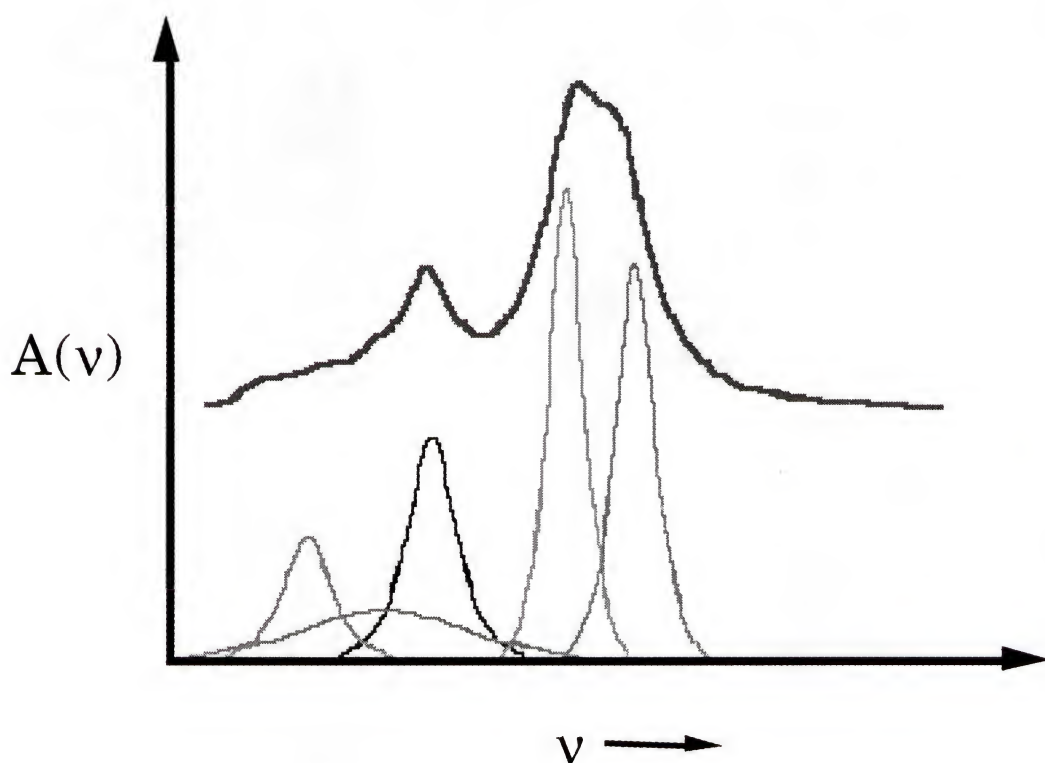


Figure 5.2: Band Overlapping

In this work we develop an empirical method for reproducing the band shape from a single geometric structure. The procedure described in the theory section of this paper is composed of two simple models. The first model approximates the width at half height for either a Gaussian or

Lorentzian band shape, while the second model allows us to estimate the oscillator strength for symmetry forbidden transitions. The Gaussian and Lorentzian band shapes are chosen because it is easily shown that low resolution spectroscopic bands possessing unresolved excitations might follow these shapes [252].

Theory

The oscillator strength, a dimensionless quantity, for the I'th transition is given by [252,253,262]

$$f_{osc}(I) = 2/3 \left| \langle \psi_0 | \mu | \psi_I \rangle \right|^2 \Delta E_{0I} \quad (\text{atomic units}) \quad 5.1a$$

$$= 4.709 \times 10^{-7} \left| \langle \psi_0 | \mu | \psi_I \rangle \right|^2 \Delta E_{0I} \quad (\text{Debye, cm}^{-1}) \quad 5.1b$$

where μ is the dipole moment operator and ΔE is the transition energy for the given transition. For comparison, the experimental oscillator strength is calculated by integrating the area under the corresponding I'th absorption peak,

$$f_{osc}(I) = 4.319 \times 10^{-9} \int \epsilon(\nu) d\nu \quad 5.2$$

where the molar extinction coefficient $\epsilon(\nu)$ is given in units of ($\text{L mol}^{-1} \text{ cm}^{-1}$), and the frequency in wavenumbers (cm^{-1}).

We have chosen to use the Gaussian functional form for the remainder of this development, although assuming a Lorentzian follows a parallel development. The functional form for the Gaussian is,

$$\varepsilon(\nu) = \bar{\varepsilon}_{\max} e^{-\alpha(\nu - \nu_{\max})^2} \quad 5.3$$

where

$$\bar{\varepsilon}_{\max} = (\varepsilon_{\max}) (4.319 \times 10^{-9}) \quad 5.4$$

and $\alpha(I)$ is

$$\alpha(I) = \frac{\pi \bar{\varepsilon}_{\max}^2(I)}{f_{osc}^2(I)} \quad 5.5$$

The calculated oscillator strength is then given by,

$$f_{osc} = \int \bar{\varepsilon}_{\max} e^{-\alpha(\nu - \nu_{\max})^2} d\nu. \quad 5.6$$

The problem now is to determine a width or a height such that the integrated area yields the calculated oscillator strength when fitted to the Gaussian form.

The width at half height for a Gaussian, equation 5.3, is

$$v_{\pm 1/2} = v_{\max} \pm \frac{0.8326}{\alpha^{1/2}} \quad 5.7a$$

or

$$= v_{\max} \pm w/2. \quad 5.7b$$

These equations are then rearranged to solve for the width (w) for each corresponding transition between the ground state (0) and excited state (I),

$$w(I) = \frac{1.6651 f_{osc}(I)}{\sqrt{\pi} \bar{\epsilon}_{\max}(I)}. \quad 5.8$$

The spectral width for a given vibrational mode between two atoms (A,B) is then estimated to be,

$$w_{AB}(I) = \langle \Delta q_{AB}^2 \rangle^{1/2} \left[\frac{\partial E(I)}{\partial q_{AB}} \right] \quad 5.9$$

In equation 5.9, $\langle \Delta q_{AB}^2 \rangle^{1/2}$ measures the motion on the ground state potential energy surface by this atom pair, while the second term measures the

change in energy of the excited state surface with respect to the change in bond distance. Assuming harmonic motion on the ground state potential energy surface, the expectation value of Δq_{AB}^2 can be derived as

$$\langle q_{AB}^2 \rangle = 2\sqrt{2} \hbar [K_{AB} \mu_{AB}]^{-1/2} \quad 5.10$$

where μ_{AB} is the reduced mass and K_{AB} ($\text{cm}^{-1} / \text{\AA}^2$) is the force constant. In general, force constants can be obtained from experiment, or calculated by diagonalizing the calculated mass weighted Hessian matrix. However, for this work, we have chosen to estimate the force constants from the AMBER force field (Table 5.1) [263].

Table 5.1: Force Constants

ATOM PAIR	Force Constant kcal/mol \AA^2	ATOM PAIR	Force Constant kcal/mol \AA^2
H-H	200.0	C-C	300.0
H-C	340.0	C-N	280.0
H-N	434.0	C-O	310.0
H-O	553.0	C-S	222.0
H-S	274.0	N-O	300.0
H-S	274.0	N-S	300.0
S-S	166.0	S-Ca	100.0
S-Sr	300.0	S-Ce	200.0

a) From Reference 263 and experimental vibrational frequencies.

The force constant for a single bond (K_{AB}^0) is used as an atom pair parameter, which is scaled according to the square root of the bond order of the appropriate atom pair.

$$K_{AB} = \sqrt{B_{AB}} K_{AB}^0 \quad 5.11$$

The use of such a scaling seems to roughly approximate the relation between single, double and triple bond K_{AB} . The bond index is defined as

$$B_{AB} = \sum_{\mu}^A \sum_{\nu}^B P_{\mu\nu} P_{\nu\mu} \quad 5.12$$

from the first-order density matrix (\mathbf{P}) formed from the symmetrical orthogonalized basis,

$$P_{\mu\nu} = \sum_a C_{\mu a} C_{\nu a} n_a \quad 5.13$$

where n_a is the occupation of orbital ϕ_i , i.e., $n_a = 0, 1$, or 2 and ϕ_i is calculated from the molecular orbital coefficients $C_{\mu a}$,

$$\phi_a = \sum_{\mu} \chi_{\mu} C_{\mu a} \quad . \quad 5.14$$

This definition of bond order (B_{AB}) has the advantage of having the traditional values of 3 for N_2 , 2 for O_2 , 1 for F_2 , etc. The use of equation 5.11 permits the use of a single empirical parameter for each pair of atoms and also has the advantage of automatically generating a zero value when there is no bond between the two atoms A and B.

In a more exact treatment of this problem one could evaluate and diagonalize the mass weighted Hessian matrix, as previously mentioned, to calculate the force constants directly for each normal mode [264-266]. We have chosen to pursue the approximate method for determining the force constants because we will be interested in very large molecules, and, in addition, the semi-empirical model Hamiltonians used to calculate the spectra of larger systems generally overestimate harmonic frequencies. Note additionally that we will sum overall atom pairs, rather than normal modes, and this will again introduce some error, as discussed below.

The derivative in equation 5.9 is only needed for the displacements specified by the ground state normal modes. This derivative could be calculated for the excited state CI wavefunction and then transformed to the ground state normal modes, but this is far from a trivial matter [267]. In the spirit of our approximations on the ground state, we have chosen to approximate this according to

$$\frac{\partial E(I)}{\partial q_{AB}} \approx \{q_{AB}(0) - q_{AB}(I)\} \times \{K_{AB}^0 \sqrt{B_{AB}(I)}\} \quad 5.15a$$

and

$$\approx Z' \times \{B_{AB}(I) - B_{AB}(0)\}^2 \times \{K_{AB}^0 \sqrt{B_{AB}(I)}\} \quad 5.15b$$

where in equation 5.15b, Z' is an empirical parameter that might be dependent upon temperature and solvent. Equation 5.15a is exact assuming a quadratic potential between A and B. The motivation for using equation 5.15b, is as follows:

1. The forces applicable on the excited state potential surface should be proportional to the change in bond strength between the ground and excited state. The change, if the potential were harmonic, would not depend on the sign of the change. To accomplish this, we take the square of the difference, although it would also be valid to take the difference to any even power or the absolute value of the difference.
2. The derivative on the excited state surface might then be proportional to $K_{AB} \sqrt{B_{AB}(I)}$, as was the case for the ground state. Figure 5.3 schematically demonstrates why proportionality to the excited state bond order, a natural part of this model, is reasonable.

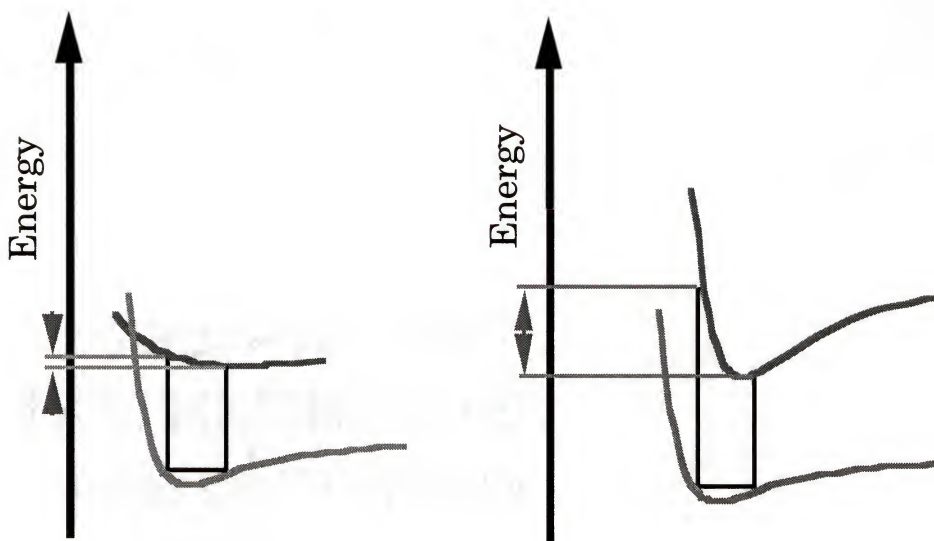


Figure 5.3: Effects of the Excited State Surface on Band Width

Equation 5.9 for each atom pair A-B is then

$$w_{AB}(I) = Z' (2\sqrt{2} \hbar)^{1/2} \frac{K_{AB}^0 \sqrt{B_{AB}(I)}}{(K_{AB}^0 \sqrt{B_{AB}(0)} \mu_{AB})^{1/4}} [B_{AB}(I) - B_{AB}(0)]^2 \quad 5.16$$

The total width for a given transition is then approximated by

$$w(I) = \sum_{A < B} w_{AB}(I) \quad 5.17$$

Strictly speaking the sum for $w(I)$ should be over the $3n-6(5)$ normal modes rather than the $n(n-1)/2$ atom pairs. We expect the errors associated with this to be small; the bond stretching (A-B) contributions are the major

contributions and are described by the neighboring atom pair interactions, the contributions from bond angles (A-B-C) are accounted for by the A-C pair interactions, the torsional or dihedral angles (A-B-C-D) interactions are accounted for by the bond stretching terms that are separated by two or more atoms. The remaining extra terms in the $n(n-1)/2$ atom pairs are very small considering the bond orders associated with these extra terms are very small.

Using the width from equation 5.17 and the calculated oscillator strength from equation 5.1, the extinction coefficient and alpha value are calculated using equations 5.6 and 5.8. The spectrum is then given by

$$A(\nu) = \sum_I \bar{\epsilon}_{\max}(I) e^{-\alpha(I) [\nu - \nu_{\max}(I)]^2} \quad 5.18$$

We would also like to account for dipole forbidden transitions that affect the appearance of the spectrum. These are especially important if the excitations are low lying and not hidden under the allowed bands. These transitions can often borrow intensity from near lying allowed transitions through vibronic coupling. This coupling is approximated by [252]

$$|\bar{\psi}_F\rangle = |\psi_F\rangle + \sum_{I \neq F} |\psi_I\rangle C_I \quad 5.19$$

$$C_I = \sum_q \frac{\langle \psi_I | \frac{\partial H}{\partial q} | \psi_F \rangle}{(E_F - E_I)} \quad 5.20$$

and

$$\langle \psi_0 | \bar{\mu} | \bar{\psi}_F \rangle = \sum_{I \neq F} \langle \psi_0 | \bar{\mu} | \psi_I \rangle C_I \quad 5.21$$

where $|\bar{\psi}_F\rangle$ is the electronically forbidden state of interest and $\langle \psi_0 | \bar{\mu} | \psi_F \rangle = 0$. $|\psi_F\rangle$ can borrow intensity from an allowed state $|\psi_I\rangle$. Equation 5.20 results from first-order perturbation theory. The oscillator strength for $|\bar{\psi}_F\rangle$ is now

$$\bar{f}_F = \frac{2}{3} |\langle \psi_0 | \bar{\mu} | \bar{\psi}_F \rangle|^2 \Delta E_{0F} \quad 5.22a$$

$$= \frac{2}{3} \left| \sum_{I \neq F} C_I \langle \psi_0 | \bar{\mu} | \psi_I \rangle \right|^2 \Delta E_{0F} \approx \sum_I \bar{f}_F(I) \quad 5.22b$$

$$\bar{f}_F(I) = \frac{C_I^2 f_I \Delta E_{0F}}{\Delta E_{0I}} \quad 5.22c$$

where we have assumed one state, usually the nearest lying allowed state, dominates in the donation of oscillator strength and we can thus drop the cross terms, as suggested in equation 5.22b. We have further assumed that nothing is lost by continuing the sum over all possible states $|\psi_I\rangle$. This approximation relieves us from having to know *a priori* which state is

lending most of the intensity. A possible side effect of this approximation is that if two or more states contribute sizably to the borrowing, there could be interference (canceling cross terms). This is not likely to be a problem because the transitions that are the most apparent through this mechanism lie at lower transition energies than any of the allowed transitions, and can therefore only borrow from those allowed transitions at higher energies.

The value for C_1 is estimated from the geometric mean.

$$C_1 = \sum_q \frac{\langle \psi_I | \frac{\partial H}{\partial q} | \psi_F \rangle}{E_F - E_I} \quad 5.23a$$

$$\approx Y \sum_q \frac{\langle \psi_I | \frac{\partial H}{\partial q} | \psi_I \rangle^{1/2} \langle \psi_F | \frac{\partial H}{\partial q} | \psi_F \rangle^{1/2}}{E_F - E_I} \quad 5.23b$$

$$\approx Y \sum_q \frac{(\partial E(I)/\partial q)^{1/2} (\partial E(F)/\partial q)^{1/2}}{E_F - E_I} \quad 5.23c$$

$$\approx Y' \sum_{A < B} K_{AB}^0 \frac{\left([B_{AB}(I) - B_{AB}(0)]^2 [B_{AB}(F) - B_{AB}(0)]^2 \right)^{1/2} [B_{AB}(I) B_{AB}(F)]^{1/4}}{E_F - E_I} \quad 5.23d$$

where $Y' = Y Z'^2$ is a second empirical constant. Note that rather cavalier approximations have been made in passing from equation 5.23a to 5.23b and

then to 5.23c. These approximations could be totally unfounded in small molecules of high symmetry. In equation 5.22b we use,

$$\bar{f}_F = \sum_{I \neq F} \bar{f}_F(I) \quad 5.24$$

If $\bar{f}_F(I) > 0.1f_I$, it is then set to $0.1f_I$. This situation has not occurred in practice, but would occur when first-order perturbation theory is no longer valid, and hence, this entire development would break down.

Method

The model outlined in the previous section was compared to experiment and to the results of molecular dynamics simulations. The experimental spectra were obtained from the Atlas of UV/Visible Spectroscopy [268]. The specifics describing the equations of motion for the molecular dynamics, and the Hamiltonian used in the calculation of the energy used in the dynamics, are developed in detail in Chapter 2 and reference 269.

The dynamics simulations were performed within the confines of the canonical ensemble (NVT). The target temperature for the dynamics was 300.0 K, with a heat bath friction constant $\tau = 0.7 \times 10^{-12}$, and a step size of 0.5×10^{-15} ps. The dynamics calculations were equilibrated until the E_{RMS} was < 0.1 kJ/mol, at which point the calculations were continued for another 50-100 ps. The coordinates used in the last 50-100 ps were outputted to a file every 0.1 ps, and these structures were subsequently used for the quantum mechanical spectroscopic calculations.

The spectroscopic calculations were all performed with the ZINDO program package [65] using the INDO/S Hamiltonian [236] and the Mataga-Nishimoto empirical two-electron integrals [270] at the CIS level of theory. The number of active orbitals included in the CIS calculations depended upon the specific system being studied, although it was always large enough to include all the active π orbitals. In some of the spectroscopic calculations, solvent effects were incorporated into the calculation using a self-consistent reaction field (SCRF) [271-273]. These SCRF calculations were performed using a single spherical cavity with radius determined from the molar mass of the solute. The SCRF interactions were calculated using the dipole and quadrupole terms.

The calculations from the molecular dynamics were plotted as a histogram of transitions, each fitted to a Lorentzian with a width at half height of 150 cm^{-1} . The empirical band width calculations used a single geometry optimized structure and the width at half height specified by equation 5.17. The following values were chosen for the empirical band width calculations: $Z' = 67.15\text{ \AA}^2$ (equation 5.15b) and $Y' = 15.00\text{ \AA}^2$ (equation 5.23d).

The "empirical band width calculation" using equations 5.18 and 5.24 required between 8 and 300 CPU seconds on an IBM RS6000 whereas the molecular dynamics simulations followed by the CIS calculation required 15 - 36 CPU hours on the same machine. The approximate model for obtaining the shape of the spectroscopic peak is thus approximately 1,000 - 10,000 times faster than the current method of using molecular dynamics to obtain band shapes, for these simple systems.

Results

Benzene and naphthalene in non-polar solvents:

The experimental UV/VIS spectrum of benzene has two symmetry forbidden bands, of ${}^1B_{2u}$ and ${}^1B_{1u}$ symmetry. These bands are easily observed experimentally, as they are in the molecular dynamics from stochastic sampling over non- D_{6h} structures. Figures 5.4 - 5.7 show the experimental spectra along with the plots obtained using molecular dynamics, approximate band width calculation, and the approximate band width calculation without estimating the oscillator strength of the forbidden bands.

The spectra obtained using the molecular dynamics is in very good agreement with experiment as seen in Figures 5.4 - 5.7. The empirical calculation for benzene should be viewed with some “skepticism”, since this was the test molecule that we used for determining the values for the two empirical parameters (Z' and Y'). Despite this caveat, the agreement between the empirical calculation, which includes both estimated band width and oscillator strength borrowing, and experiment or molecular dynamics is very good. The numerical values used for creating the spectra in Figures 5.4 - 5.7 are listed in Table 5.2. The experimental and calculated spectra for naphthalene are all given in Figures 5.8 - 5.11. Again, as in benzene, the symmetry forbidden bands have an important role in determining the appearance of the spectrum. Comparison of the two oscillator strengths are presented in Table 5.3.

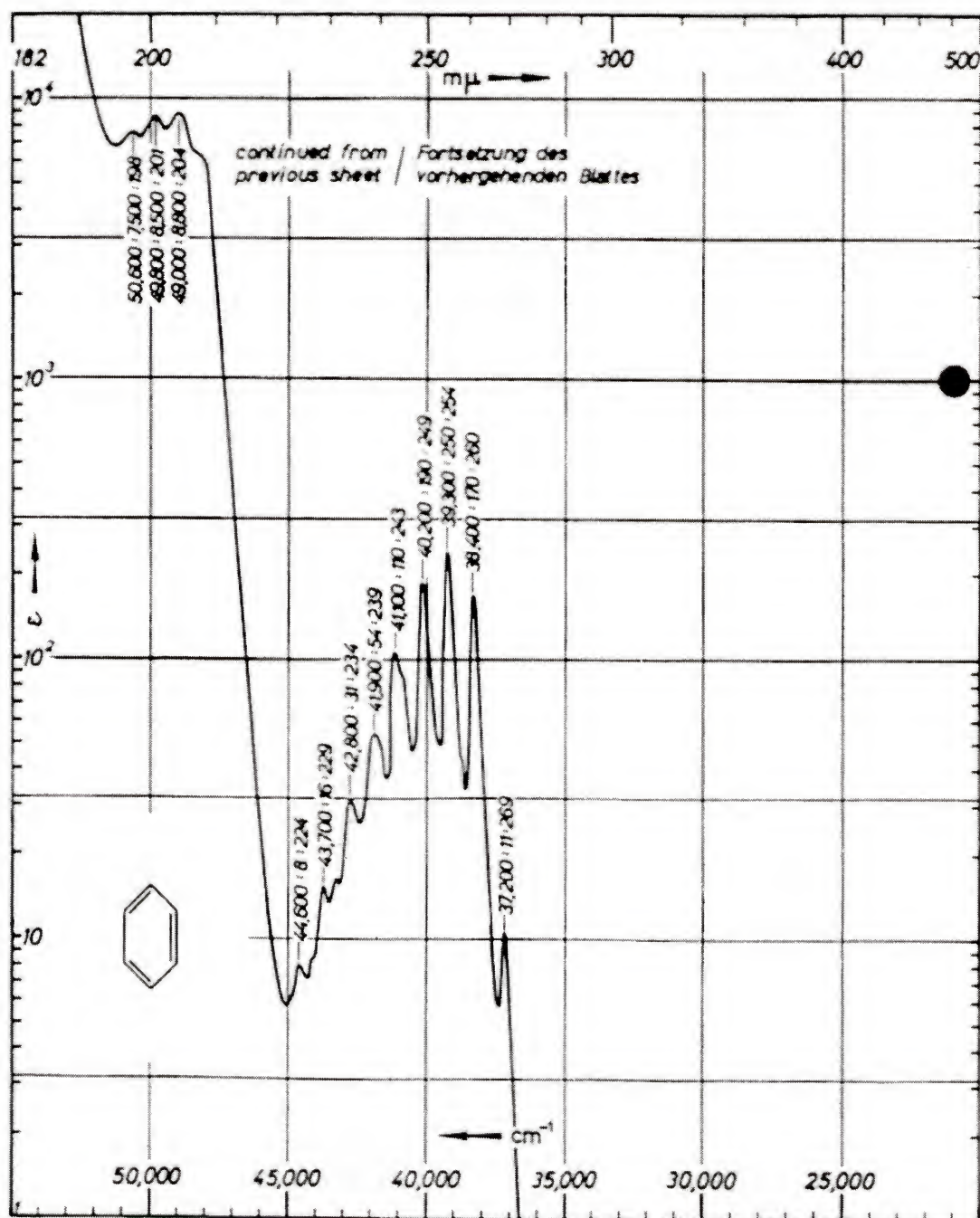


Figure 5.4: Experimental UV Spectra of Benzene

Reference 269

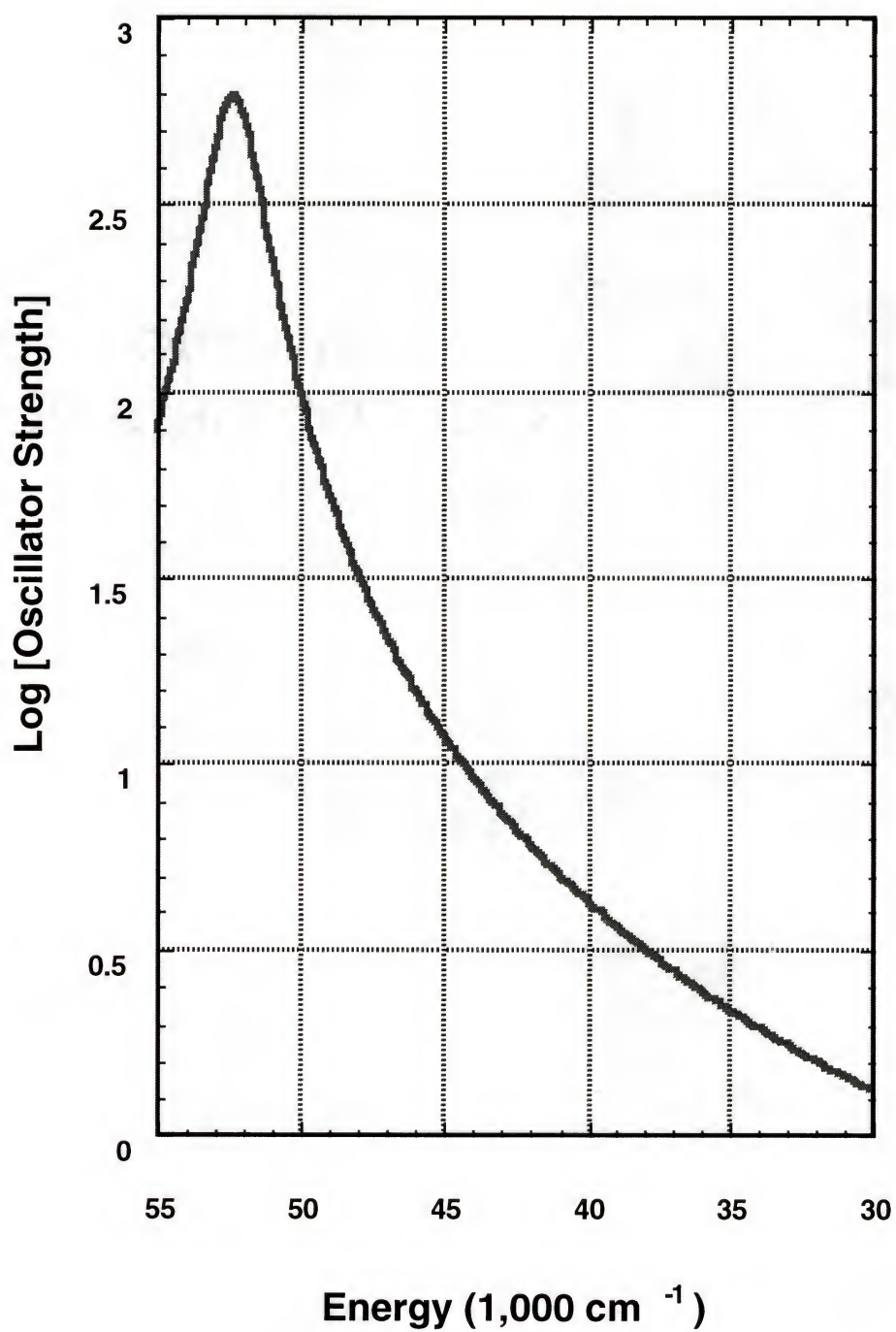


Figure 5.5: Calculated UV Spectra of Benzene Using the Band Width Theory

Equation 5.15b

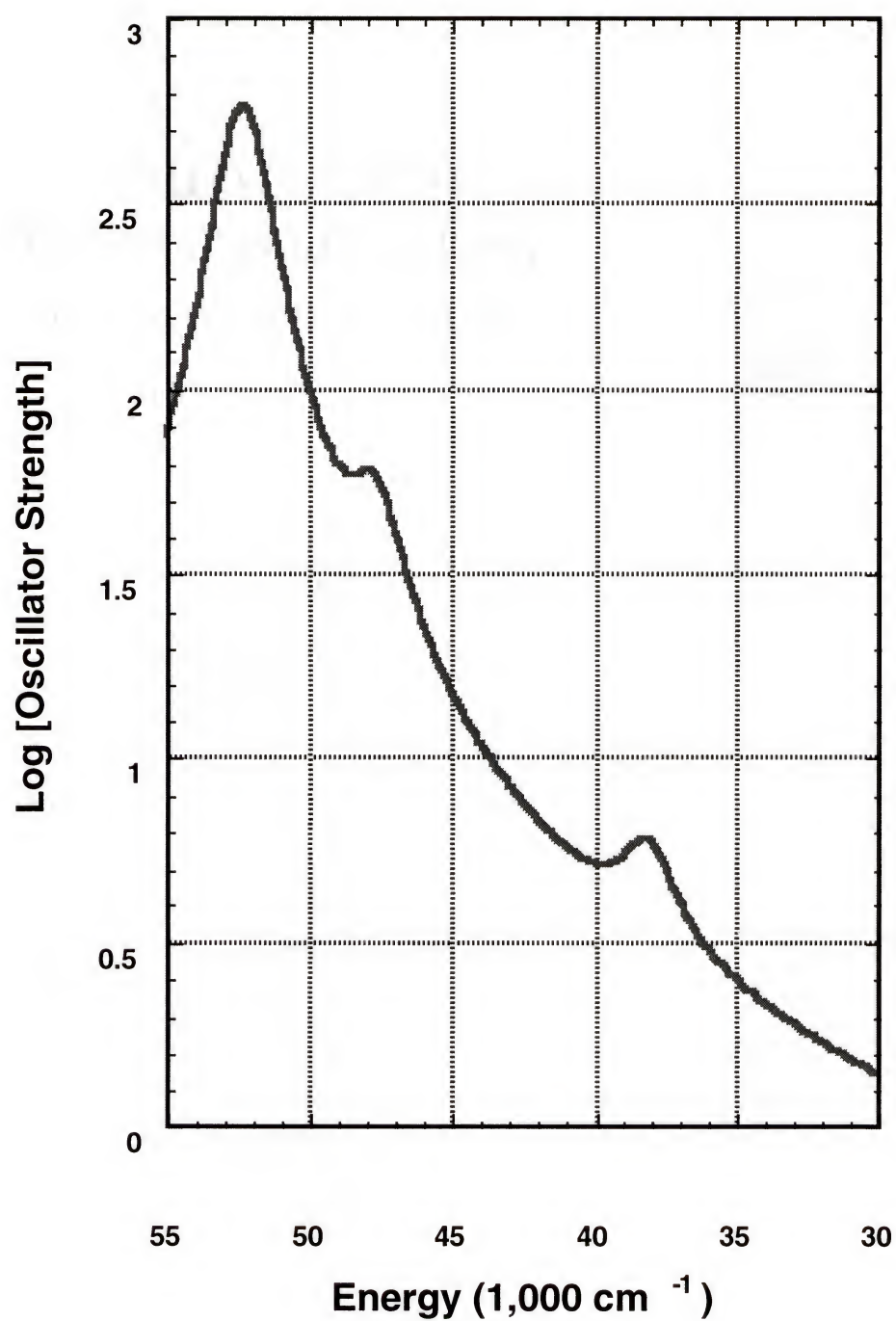


Figure 5.6: Calculated UV Spectra of Benzene Using the Band Width Theory and Transition Strength Correction

Equations 5.15b and 5.23d

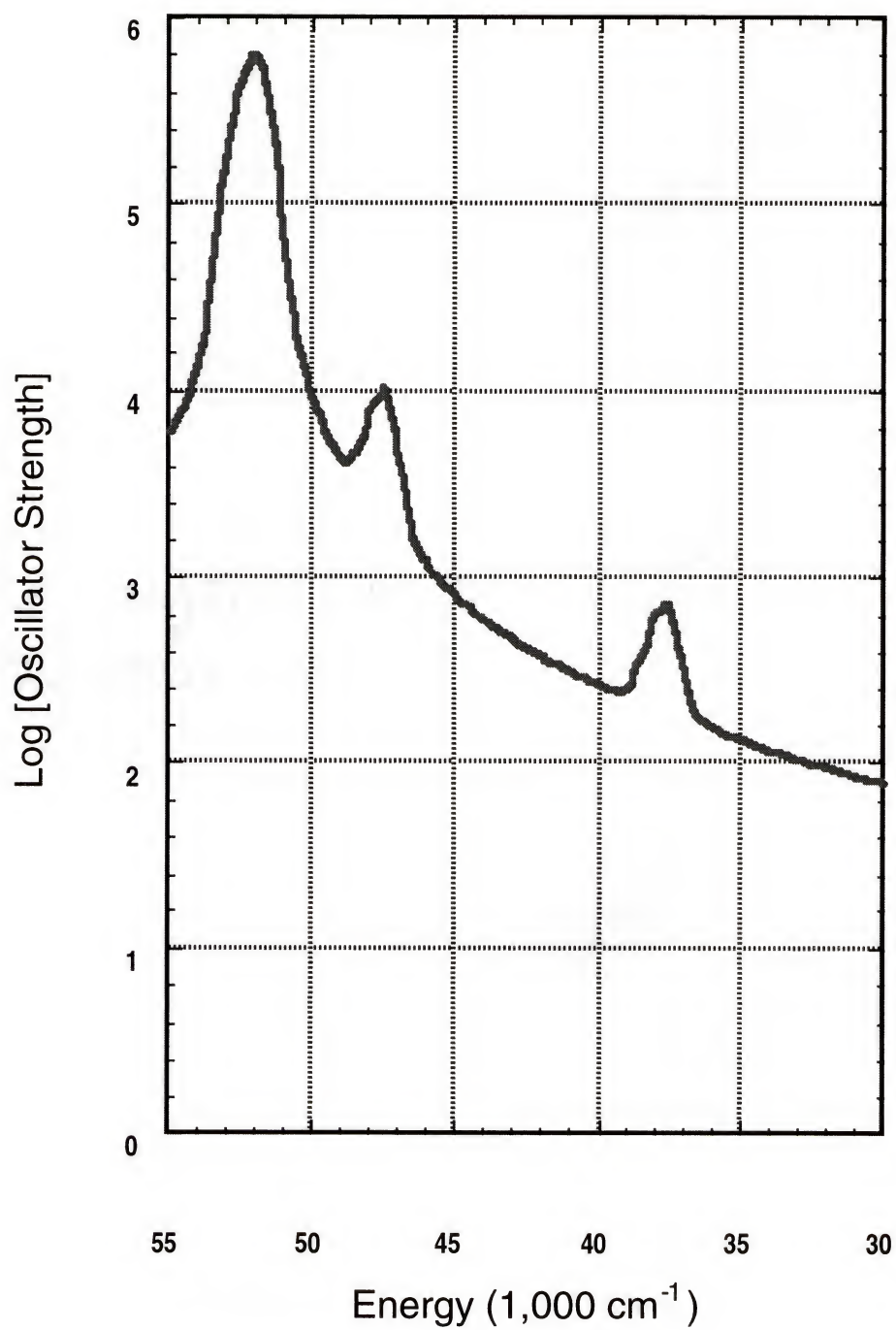


Figure 5.7: Calculated UV Spectra of Benzene From Molecular Dynamics Simulation

Table 5.2: Calculated Benzene Transitions

States	Energy (1,000 cm ⁻¹)	Experiment (1,000 cm ⁻¹)	Width (1,000 cm ⁻¹)	Osc Str Calc ^a	Osc Str Est ^a
¹ B _{2u}	38.1	~40	2.0	0.00	0.01
¹ B _{1u}	47.7	~50	2.0	0.00	0.10
¹ E _u	52.3 (x2)	~54	2.1	0.99	0.94

- a) The calculated oscillator strengths are direct from the quantum mechanics, Equation 5.1. The estimated oscillator strengths use the calculated oscillator strengths and the approximation of Equation 5.23.

Table 5.3: Calculated Naphthalene Transitions

State	Energy (1,000 cm ⁻¹)	Experiment (1,000 cm ⁻¹)	Width (1,000 cm ⁻¹)	Osc Str calc ^a	Osc Str est ^a
¹ B _{3u}	32.0	~32	1.4	0.004	0.004
¹ B _{2u}	35.1	~35 - 38	3.0	0.18	0.17
¹ B _{3u}	43.2		1.3	1.74	1.31
¹ B _{1g}	44.2		3.5	0.00	0.24
¹ A _g	44.5	~45	1.5	0.00	0.23
¹ B _{2u}	46.1		1.6	0.63	0.49
¹ B _{2g}	49.0		2.0	0.00	0.04
¹ B _{1g}	49.9		3.5	0.00	0.06
¹ A _g	53.7		2.7	0.00	0.02

- a) The calculated oscillator strengths are direct from the quantum mechanics, Equation 5.1. The estimated oscillator strengths use the calculated oscillator strengths and the approximation of Equation 5.23.

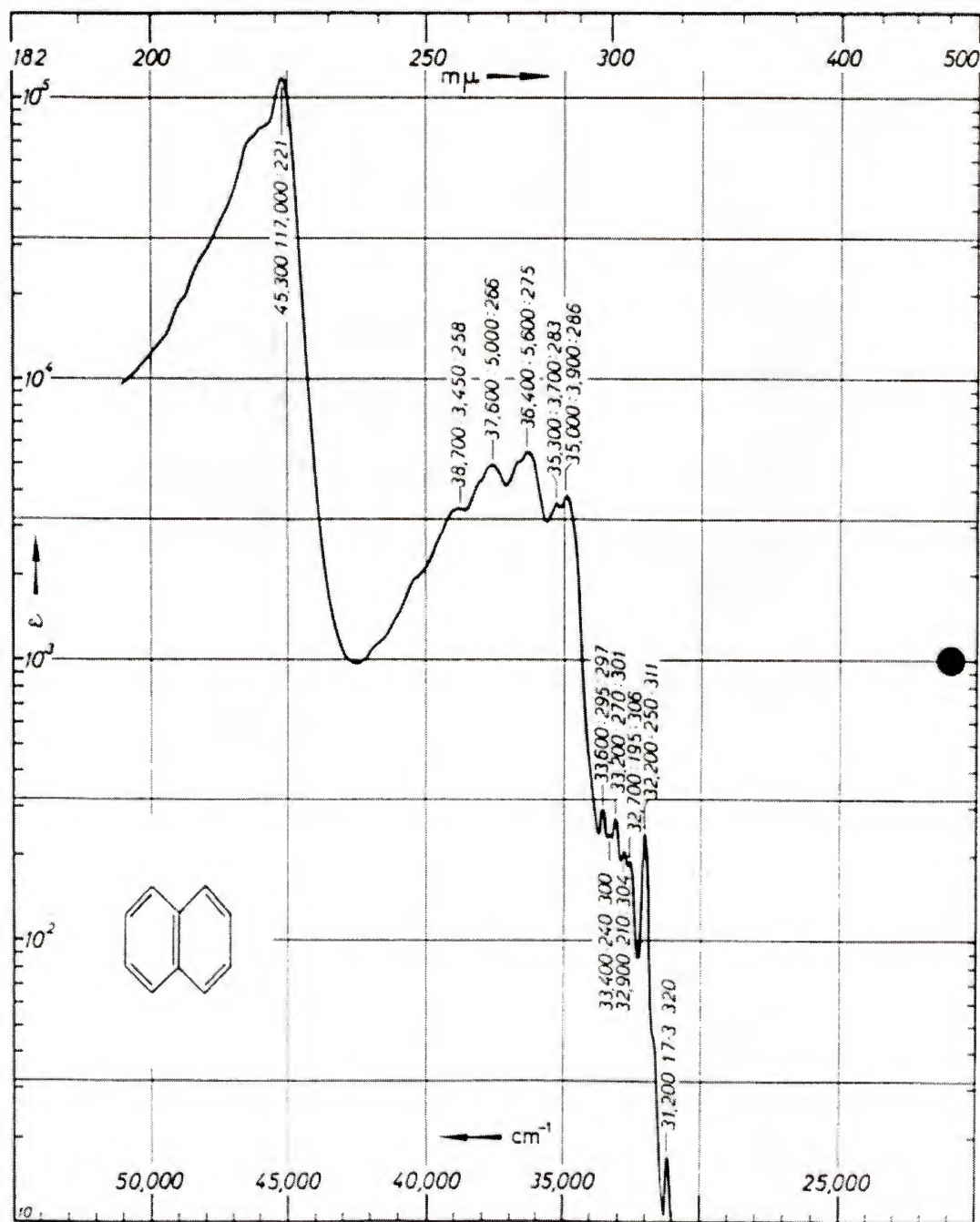


Figure 5.8: Experimental UV Spectra of Naphthalene

Reference 269

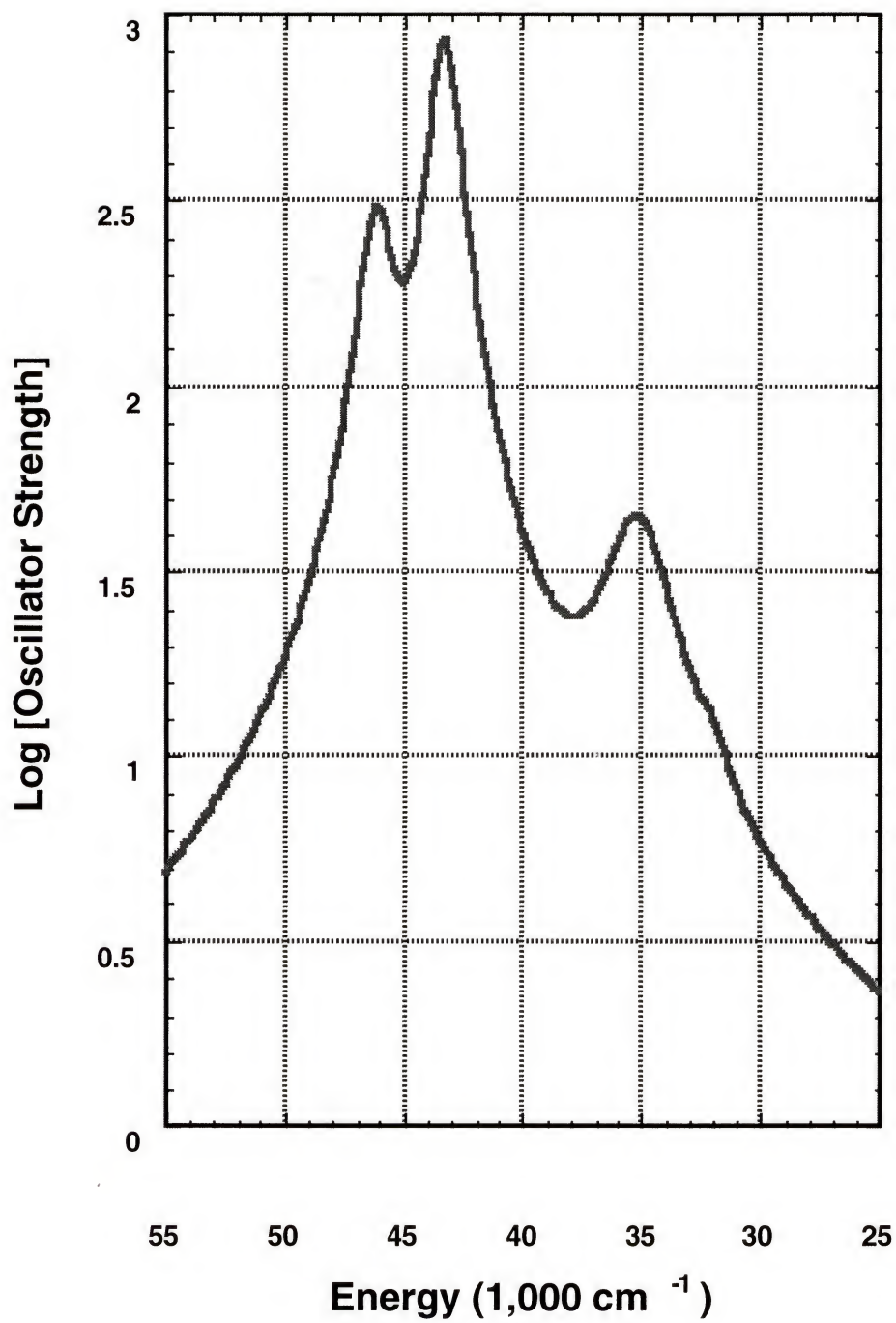


Figure 5.9: Calculated UV Spectra of Naphthalene Using the Band Width Theory

Equation 5.15b

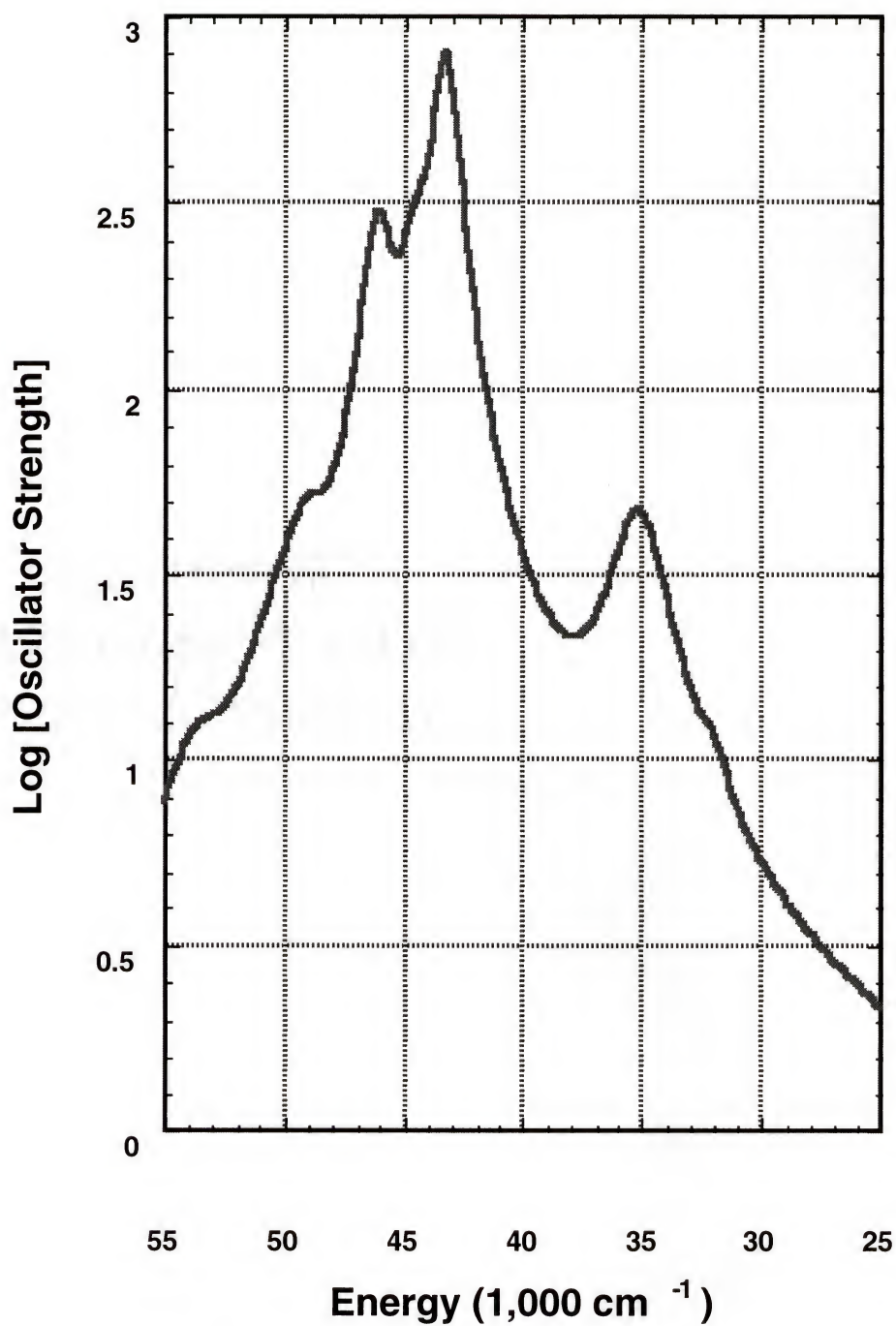


Figure 5.10: Calculated UV Spectra of Naphthalene Using the Band Width Theory and Transition Strength Correction

Equations 5.15b and 5.23d

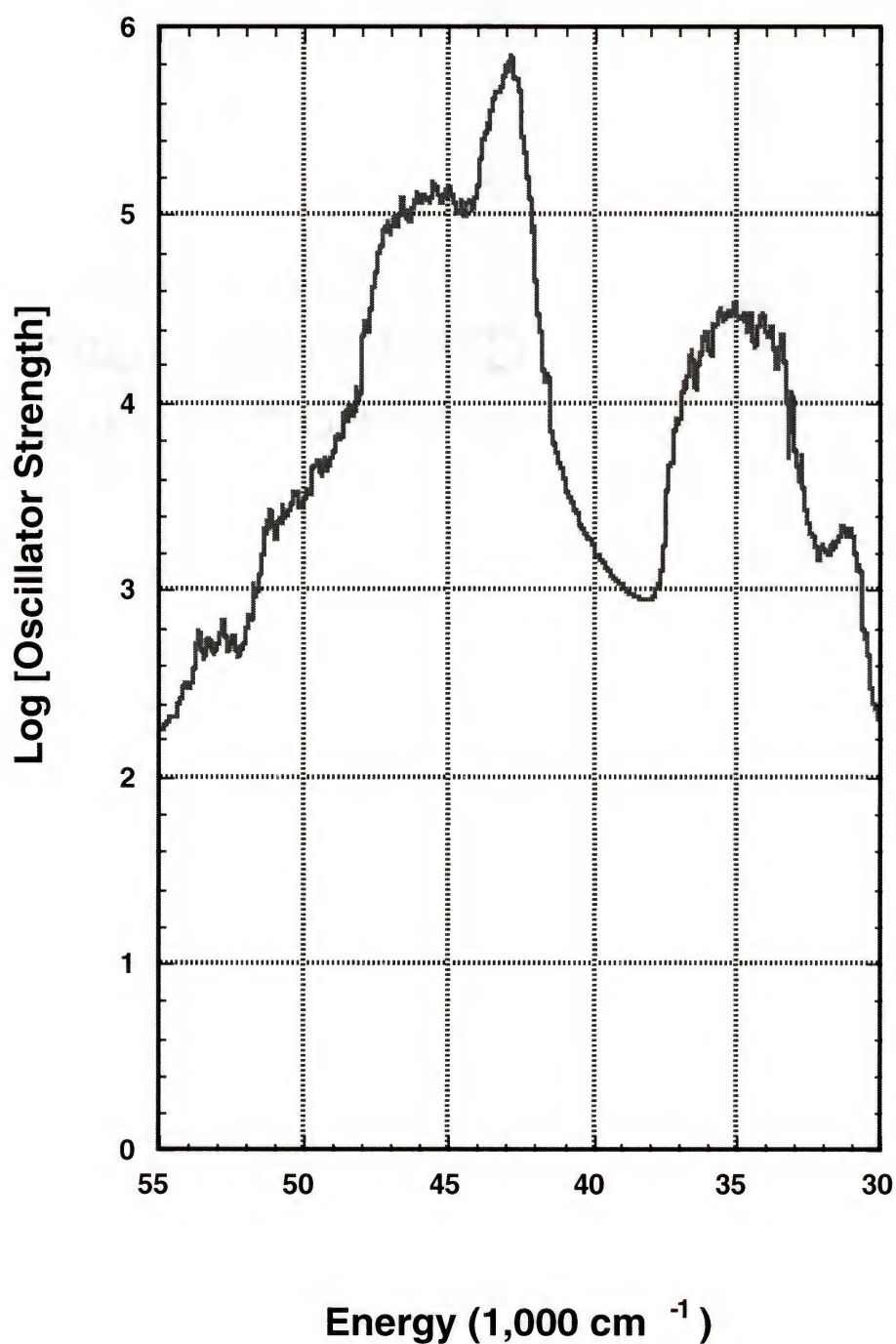


Figure 5.11: Calculated UV Spectra of Naphthalene From Molecular Dynamics Simulation

Comparison of different solvent for pyridine, pyrimidine, and pyridazine:

The previous sections showed that the empirical band width calculations are capable of modeling the appearance of simple organic molecules in non-polar solvents. We decided to determine if the single empirical parameter Z' is able reproduce several different solvents, or if each solvent requires a significantly different constant. We have chosen to study pyridine, pyrimidine, and pyridazine in order to test the solvent dependence of Z' . These molecules were chosen because they have a permanent dipole and the appearance of the spectrum is known to be somewhat sensitive to different solvents. The non-polar solvent is modeled with a cyclohexane reaction field. The polar solvent is modeled with a water reaction field, whereas the polar + $n(\text{H}_2\text{O})$ is modeled with a reaction field and n water molecules at hydrogen bond distances determined by the quantum chemistry.

The experimental spectrum of pyridine shows that the first observed peak is narrower in water when compared to cyclohexane. This is a consequence of the $n-\pi^*$ excitation being blue shifted under the first allowed $\pi-\pi^*$ transition. This phenomenon is reproduced in the empirical band width calculation. There is a good agreement between the appearance of the experimental and calculated spectra (see Figures 5.12 - 5.17). The calculated spectra have also reproduced the deeper minimum between the first and second peak as observed in the experimental spectra. In Table 5.4, the calculated peaks along with their corresponding widths at half height are presented for several different environments.

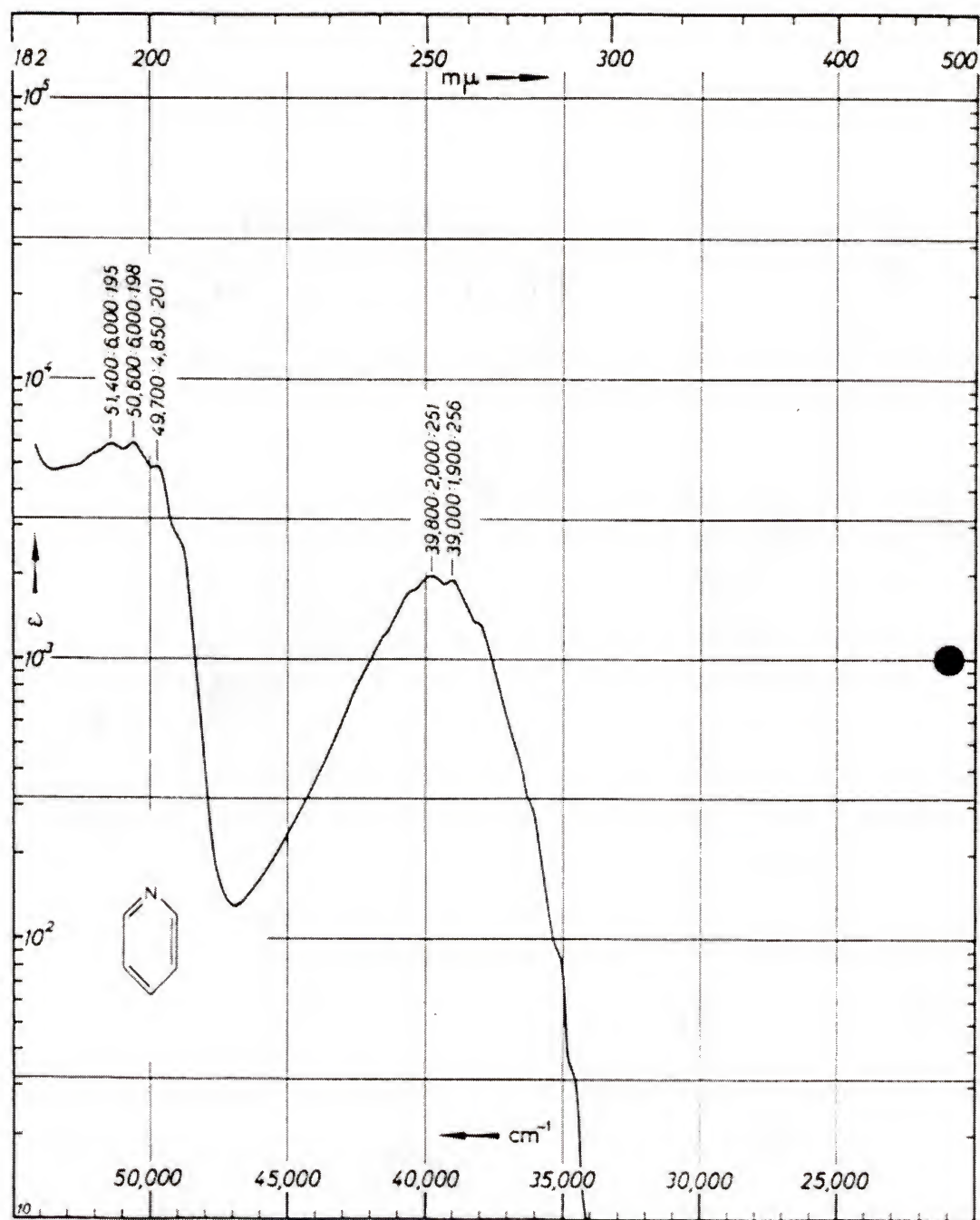


Figure 5.12: Experimental UV Spectra of Pyridine in Non-polar Solvent

Reference 269

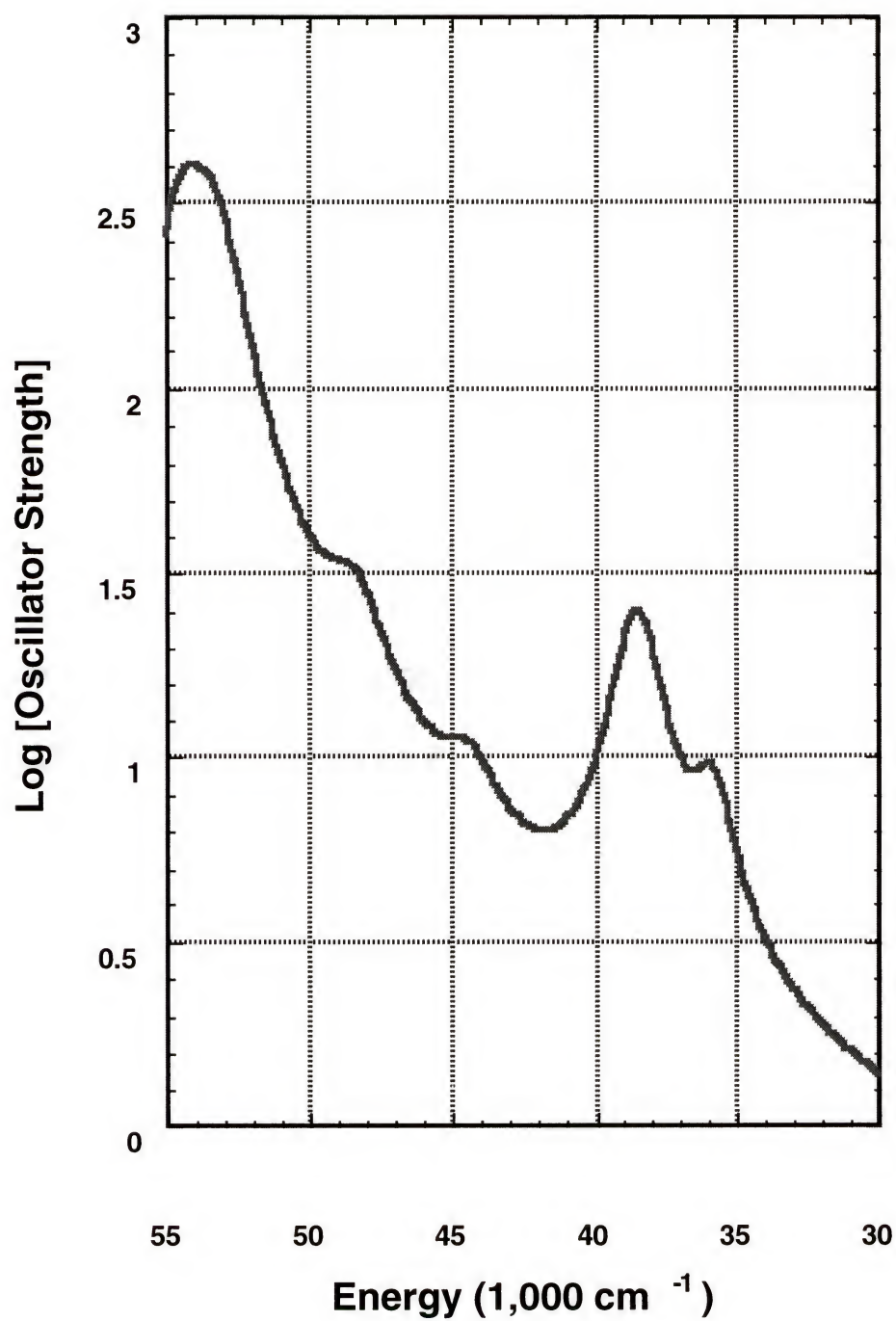


Figure 5.13: Calculated UV Spectra of Gas Phase Pyridine Using the Band Width Theory

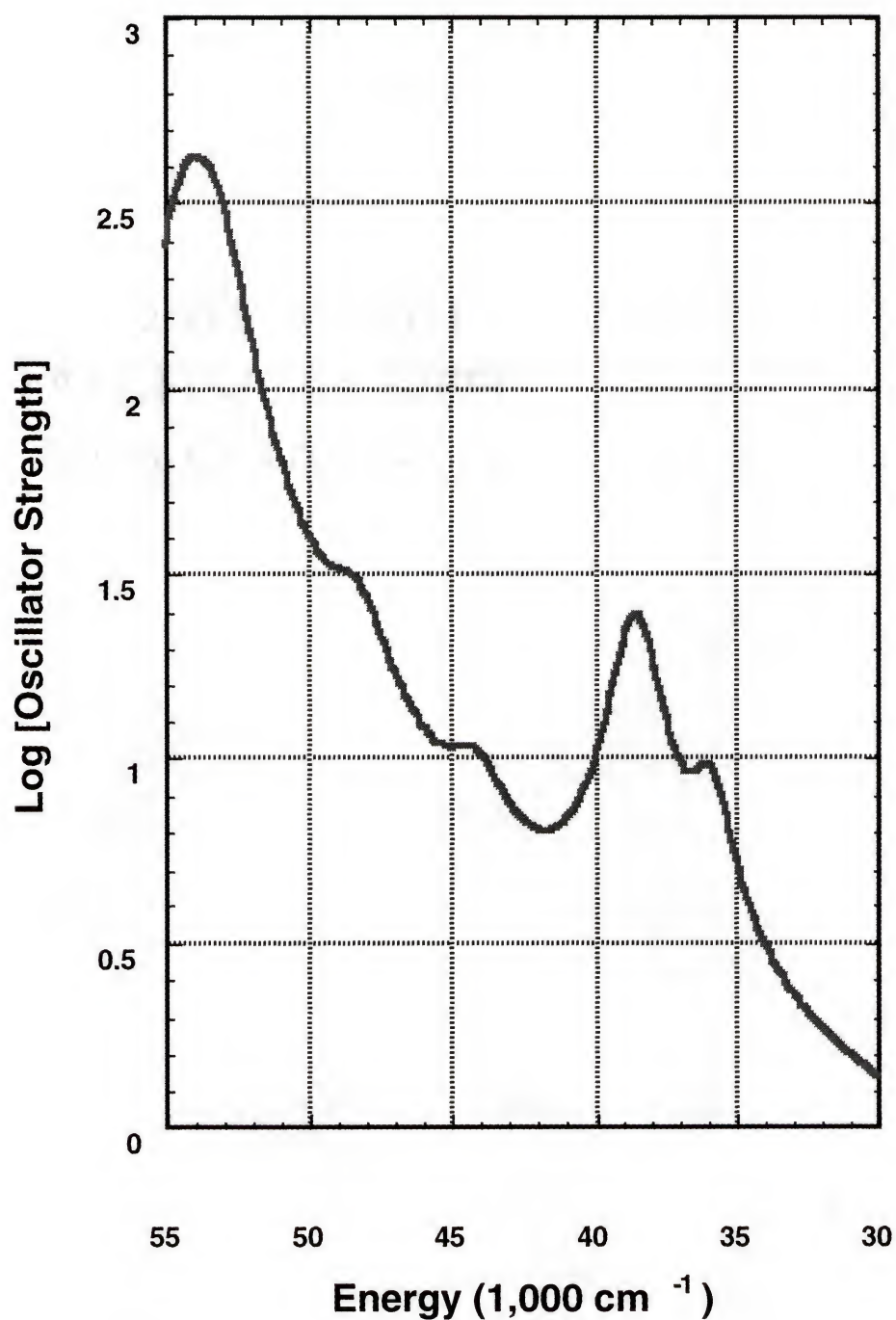


Figure 5.14: Calculated UV Spectra of Pyridine in Non-polar Solvent Using the Band Width Theory

Spherical cavity SCRF calculation.

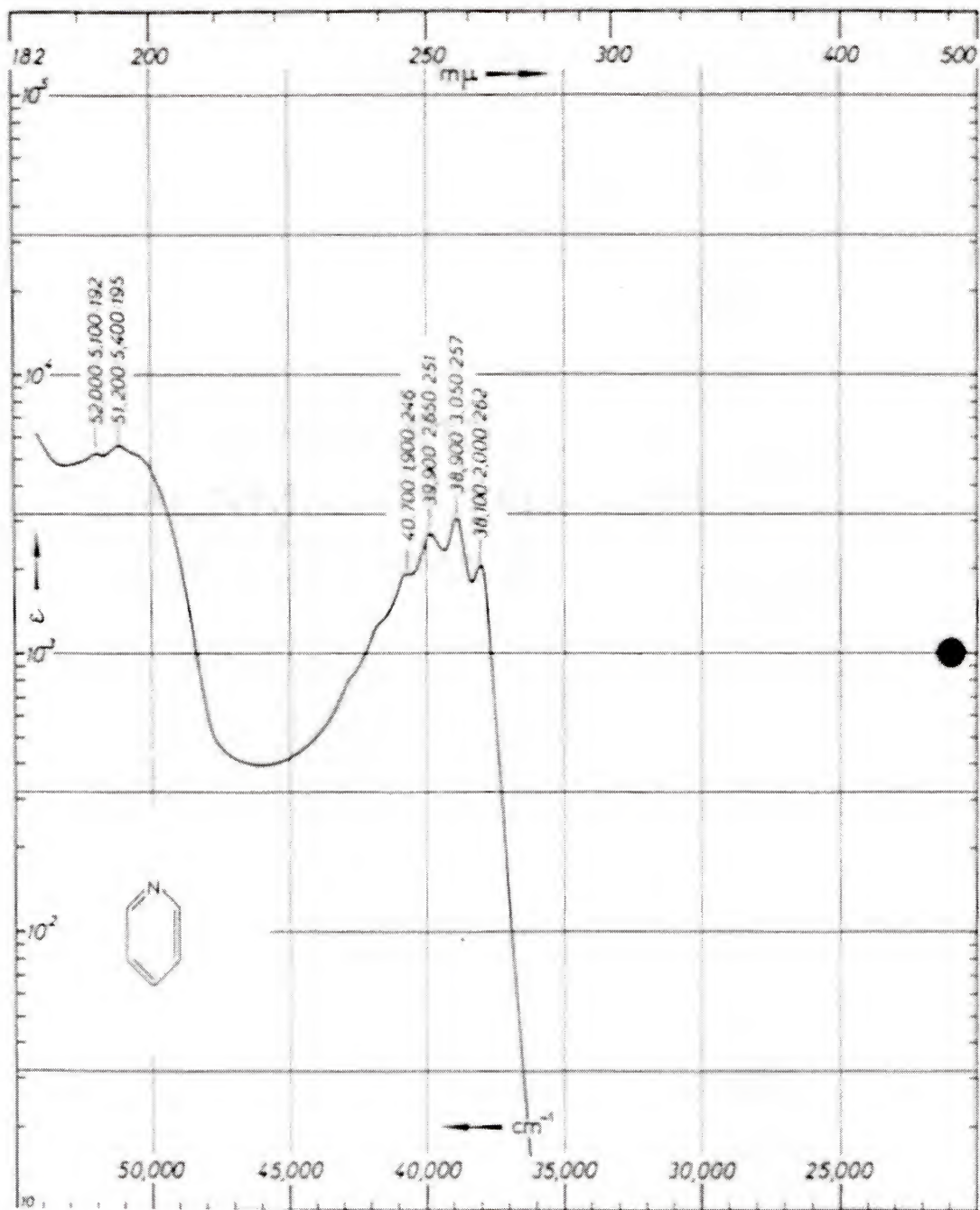


Figure 5.15: Experimental UV Spectra of Pyridine in Water

Reference 269

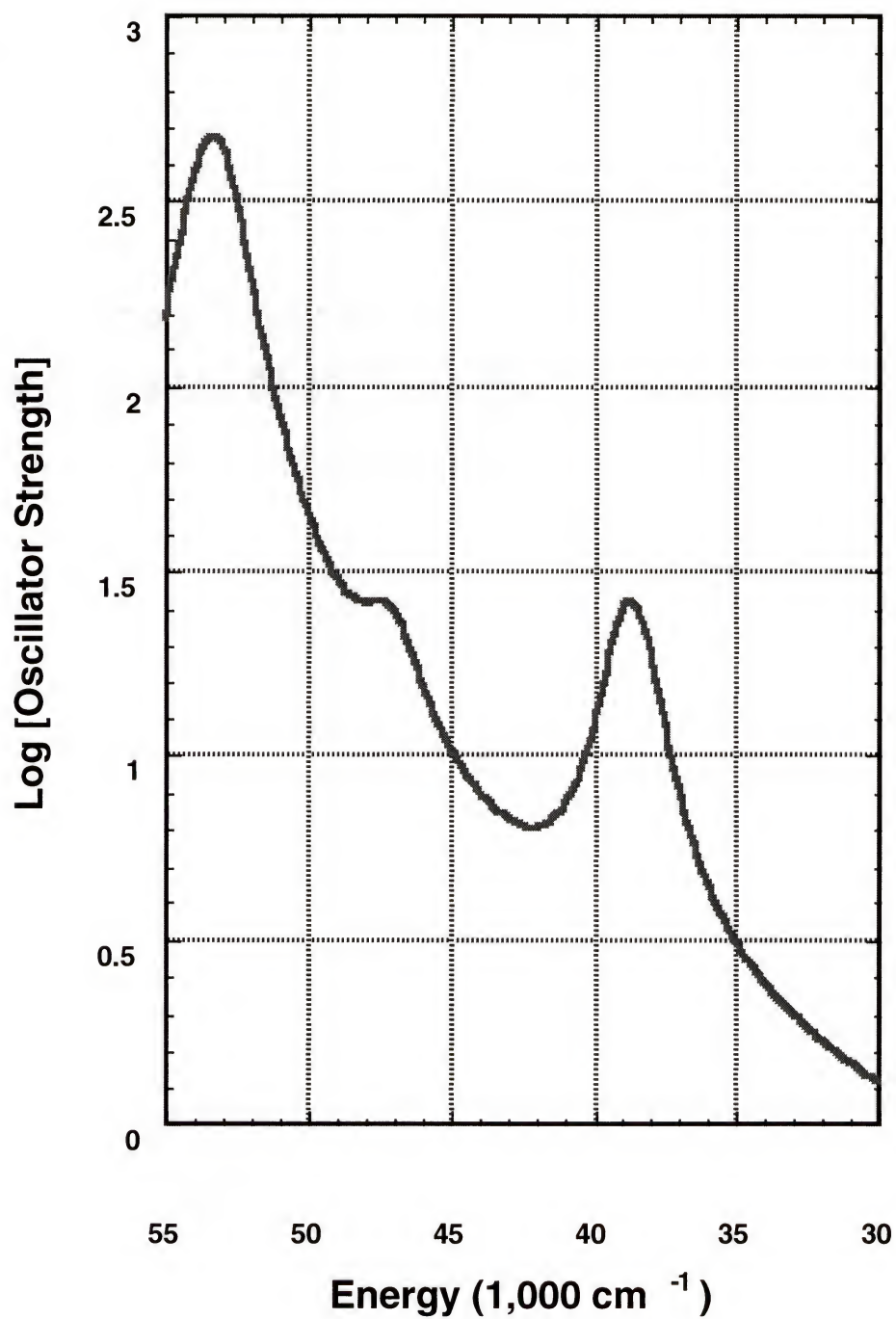


Figure 5.16: Calculated UV Spectra of Pyridine Water Using the Band Width Theory

Spherical cavity SCRF calculation.

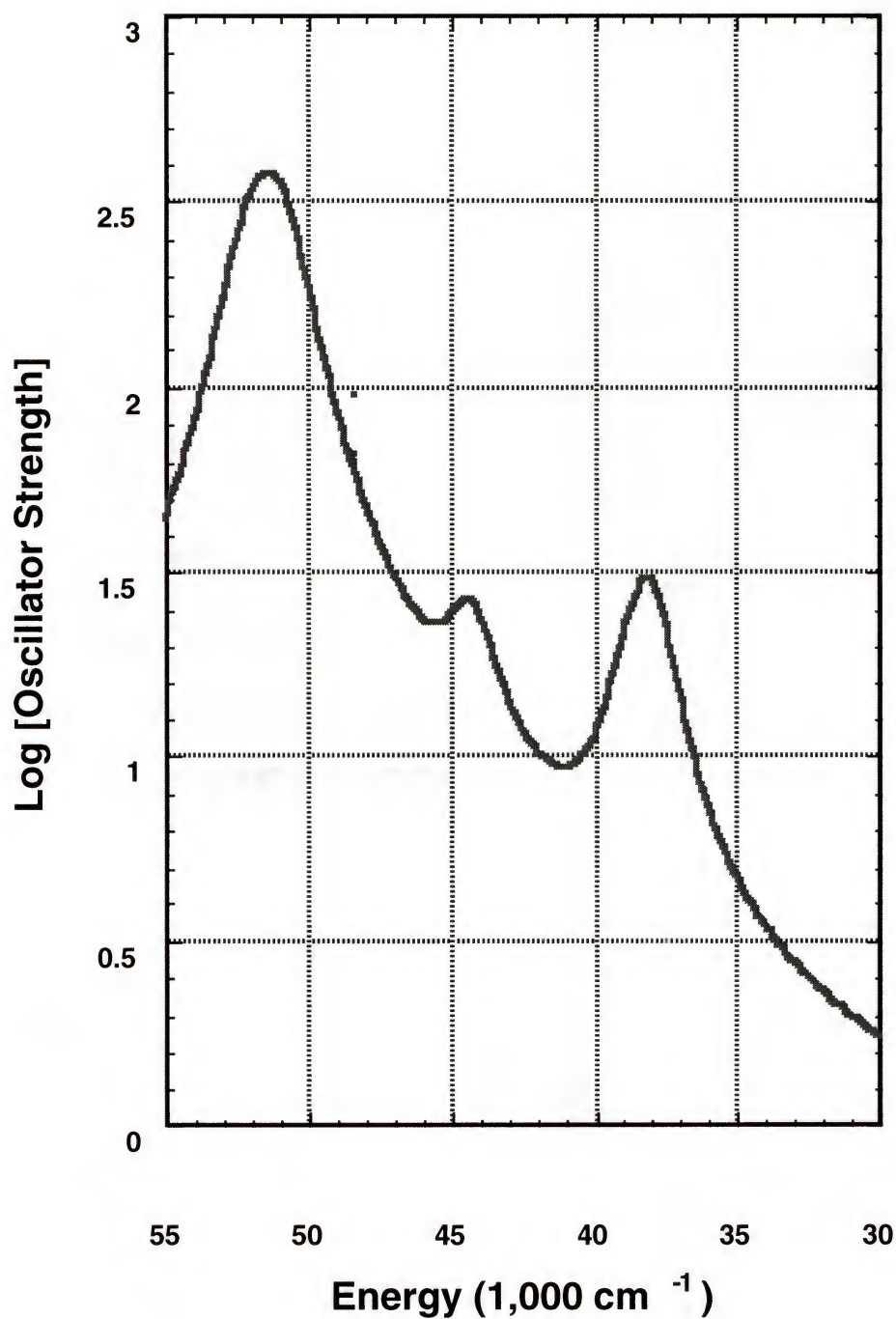


Figure 5.17: Calculated UV Spectra of Pyridine Water Using the Band Width Theory 2

Spherical cavity SCRF calculation on supermolecule structure, nearest hydrogen bonding water included in the QM calculation.

Table 5.4: Comparison of Different Solvents for Pyridine
All numbers listed in $1,000\text{ cm}^{-1}$ ^a

Exp.		Gas Phase		Non-polar		Polar		Polar + H ₂ O	
		E	Width	E	Width	E	Width	E	Width
¹ B ₂	34.7	35.8	1.6	35.9	1.6	38.1	1.5	36.9	1.6
¹ B ₁	38.5	38.5	1.9	38.5	1.9	38.8	1.9	38.0	1.9
¹ A ₂		44.4	2.0	44.1	2.1	47.1	2.1	44.3	2.0
¹ A ₁	49.7	48.4	2.3	48.3	2.3	47.8	2.3	48.4	2.5
¹ A ₁	55.0	53.3	2.2	53.3	2.1	53.0	2.0	51.0	2.3
¹ B ₁	56.4	54.4	2.0	54.3	2.0	53.7	2.0	51.9	2.3

- a) The symmetry labels were calculated from the C_{2v} point group by ZINDO for all of the above calculations except for the reaction field + explicit water.

The gas phase entry are the results obtained without using a reaction field or specific modeling of the solvent. The non-polar column is obtained from a reaction field modeling cyclohexane for the solvent, $\epsilon = 2.023$, $\eta = 1.4266$. The polar solvent is also a reaction field modeling water for the solvent, $\epsilon = 78.50$, $\eta = 1.3332$. Finally, the polar + H₂O uses the water SCRF and adds either 1 or 2 explicit water molecules. These water molecules are placed in hydrogen bonding positions with respect to the solute as determined by the simulations [274].

The pyrimidine spectra are also well reproduced as seen in Figures 5.18 - 5.23. The first peak, as in pyridine, is narrower in the water spectra when compared to the non-polar solvent.

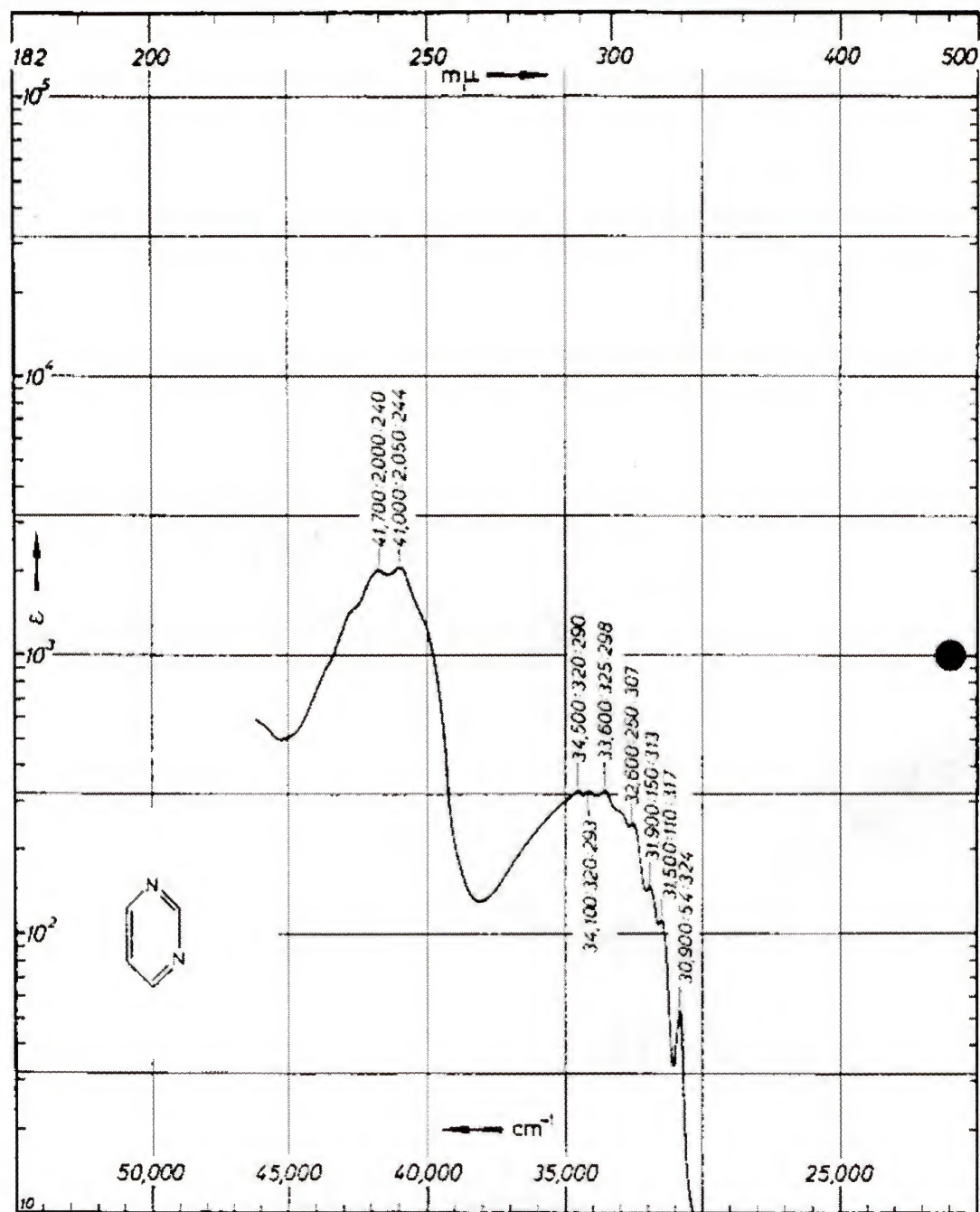


Figure 5.18: Experimental UV Spectra of Pyrimidine in Non-polar Solvent

Reference 269

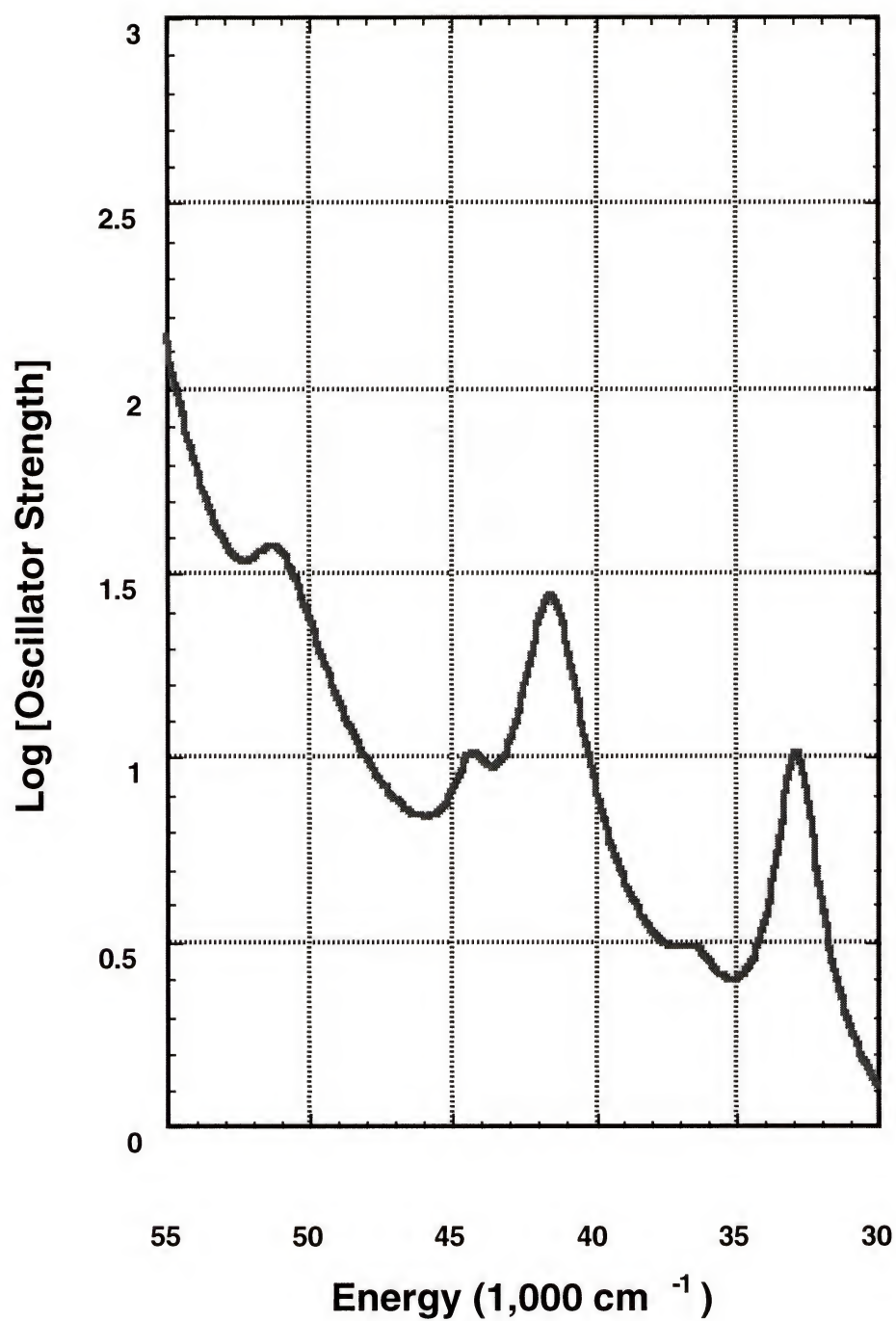


Figure 5.19: Calculated UV Spectra of Gas Phase Pyrimidine Using the Band Width Theory

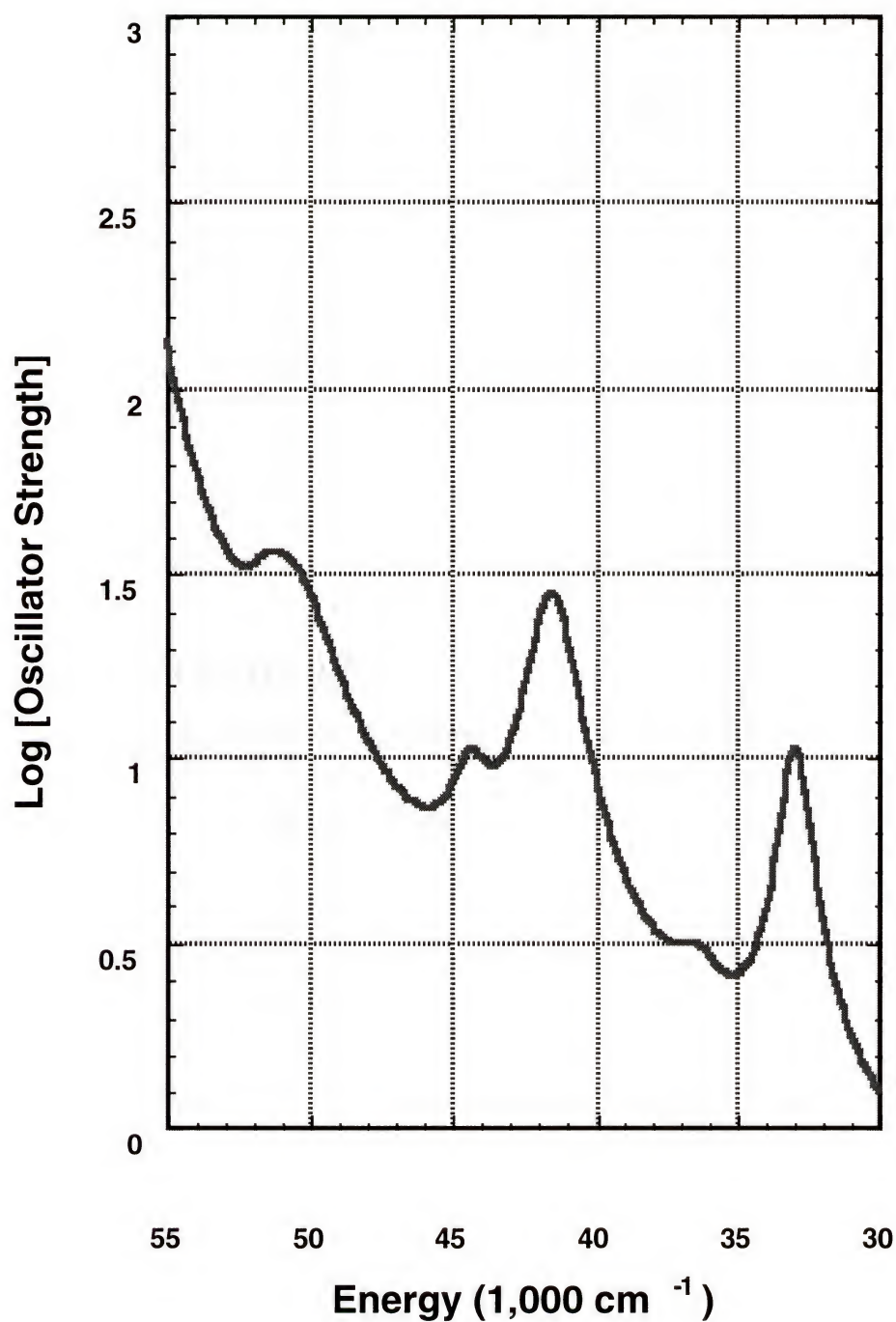


Figure 5.20: Calculated UV Spectra of Pyrimidine in Non-polar Solvent Using the Band Width Theory

Spherical cavity SCRF calculation.

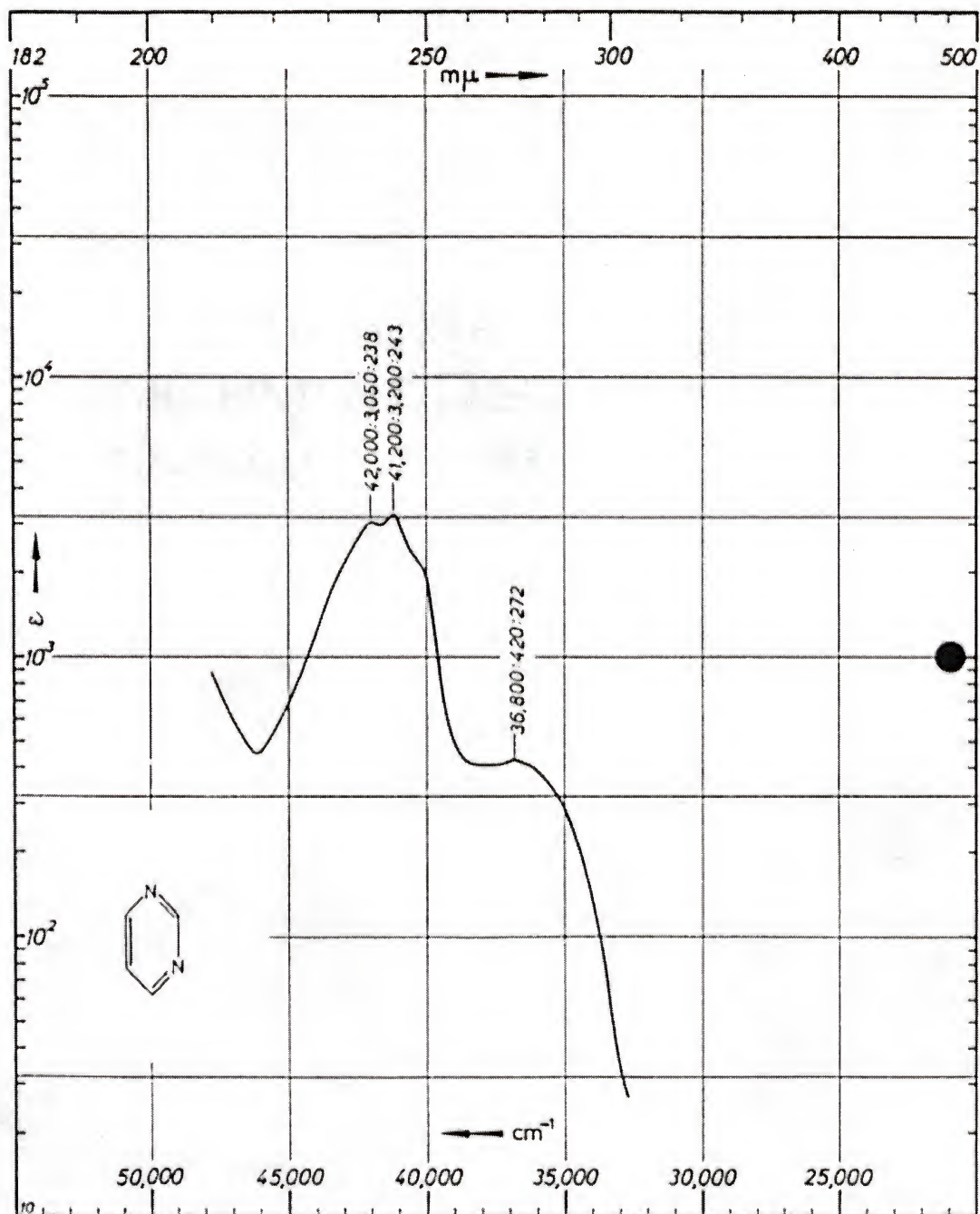


Figure 5.21: Experimental UV Spectra of Pyrimidine in Water

Reference 269

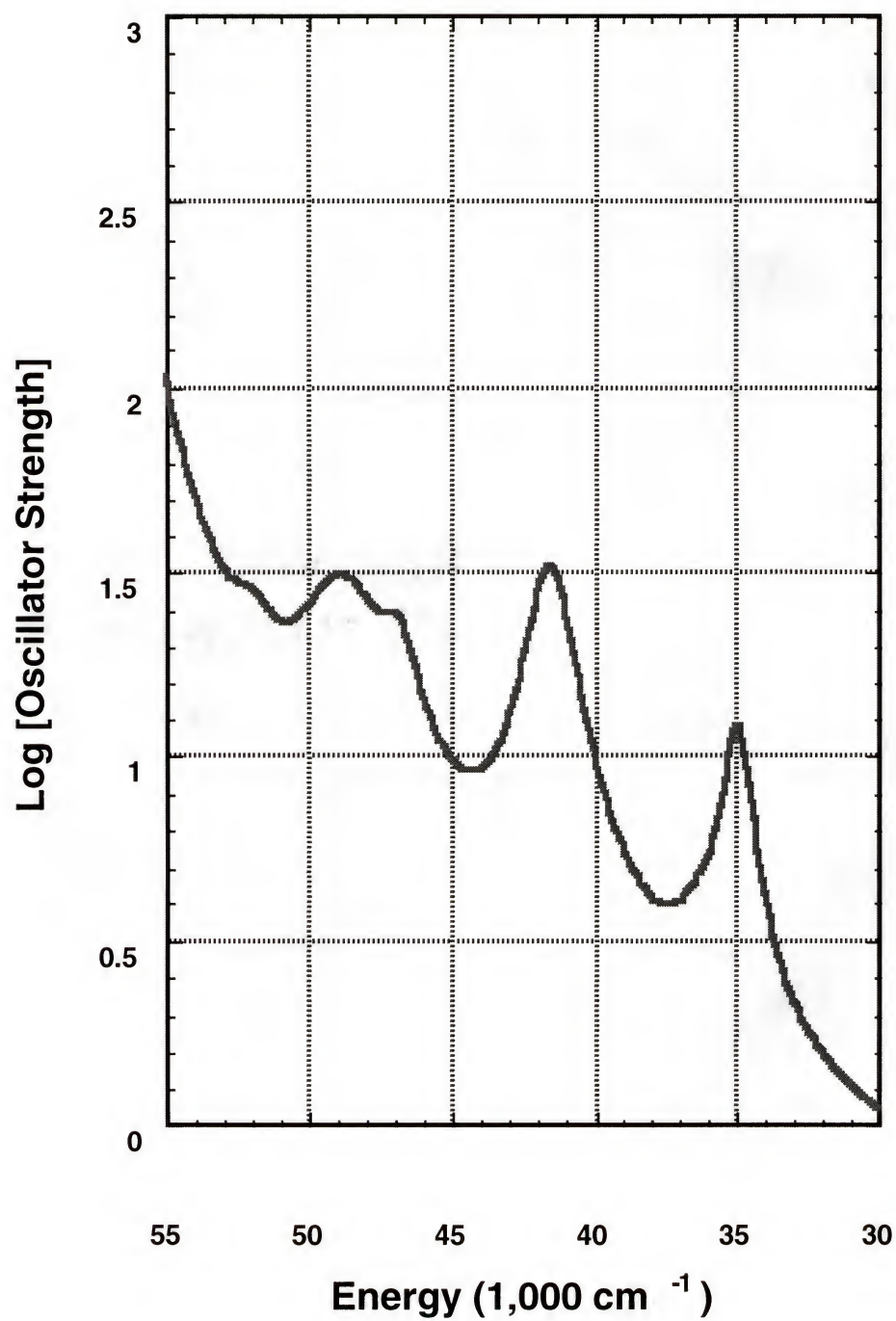


Figure 5.22: Calculated UV Spectra of Pyrimidine Water Using the Band Width Theory

Spherical cavity SCRF calculation.

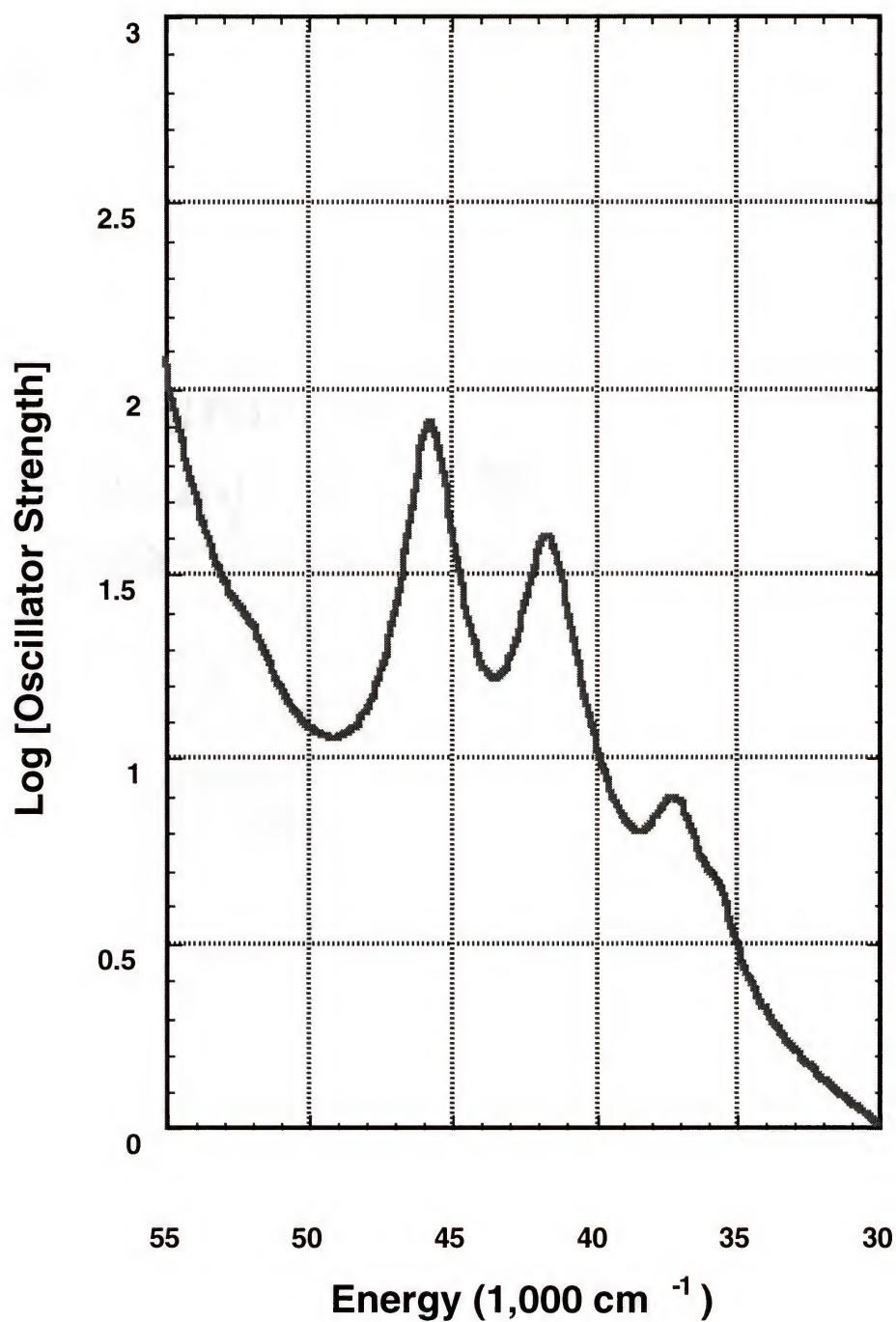


Figure 5.23: Calculated UV Spectra of Pyrimidine Water Using the Band Width Theory 2

Spherical cavity SCRF calculation on supermolecule structure, nearest hydrogen bonding water included in the QM calculation.

This trend is again duplicated by the empirical band width calculations, which can be observed in Figures 5.18 - 5.23 and Table 5.5. Note the rather large differences in calculated widths between polar calculations and the calculations that actually models explicitly bound water. This is a good indicator that a single Z' parameter should be sufficient for modeling all solvents.

Table 5.5: Comparison of Different Solvents for Pyrimidine
All numbers listed in $1,000\text{ cm}^{-1}$ ^a

Exp.		Gas Phase		Non-polar		Polar		Polar + 2 H ₂ O	
		E	Width	E	Width	E	Width	E	Width
¹ B ₂	31.0	32.9	1.2	33.0	1.2	34.9	1.1	35.6	1.1
¹ A ₂		36.4	1.8	36.4	1.9	36.2	2.0	37.1	2.0
¹ B ₁	40.3	41.5	1.8	41.5	1.8	41.5	1.8	41.6	1.8
¹ A ₂		44.3	1.5	44.3	1.4	46.8	1.4	45.7	1.5
¹ A ₁	51.1	50.9	2.8	50.6	2.9	48.8	3.4	49.1	3.5
¹ B ₂	52.3	51.3	1.5	51.5	1.5	51.9	1.5	52.0	1.8
¹ A ₁	56.2	56.2	2.0	56.3	2.0	56.7	1.9	56.2	2.0
¹ B ₁	58.5	56.9	1.9	56.9	1.9	56.7	2.0	57.3	2.0

- a) The symmetry labels were calculated from the C_{2v} point group by ZINDO for all of the above calculations except for the reaction field + explicit water.

The spectrum of pyridazine is also well reproduced using the empirical band width approximation (see Figures 5.24 - 5.28 and Table 5.6). While the spectra calculated for these three molecules correctly models changes to the band shape, the calculated changes are not as pronounced as they are in the experimental spectra. We are currently parameterizing Z' to better reproduce these differences.

Table 5.6: Comparison of Different Solvents for Pyridazine
All numbers listed in $1,000\text{ cm}^{-1}$ ^a

Exp.		Gas Phase		Non-polar		Polar		Polar + 2 H ₂ O	
		E	Width	E	Width	E	Width	E	Width
¹ B ₁	26.6	29.6	1.5	30.2	1.5	33.5	1.8	33.7	1.9
¹ A ₂	39.5	37.4	1.6	37.9	1.6	42.1	1.6	37.6	1.7
¹ A ₁		37.5	1.9	37.7	1.9	38.7	1.9	39.7	1.9
¹ A ₂		44.4	1.5	44.8	1.5	48.6	1.4	43.3	1.5
¹ B ₂		46.6	2.0	46.3	2.0	43.5	2.0	49.7	1.9
¹ B ₂	~50	52.1	2.0	51.9	2.0	51.8	2.0	50.2	1.9
¹ B ₁		52.1	1.7	52.4	1.7	57.3	1.8	52.8	3.1
¹ A ₁		52.4	2.0	52.3	2.0	51.8	2.6	54.3	3.3

- a) The symmetry labels were calculated from the C_{2v} point group by ZINDO for all of the above calculations except for the reaction field + explicit water.

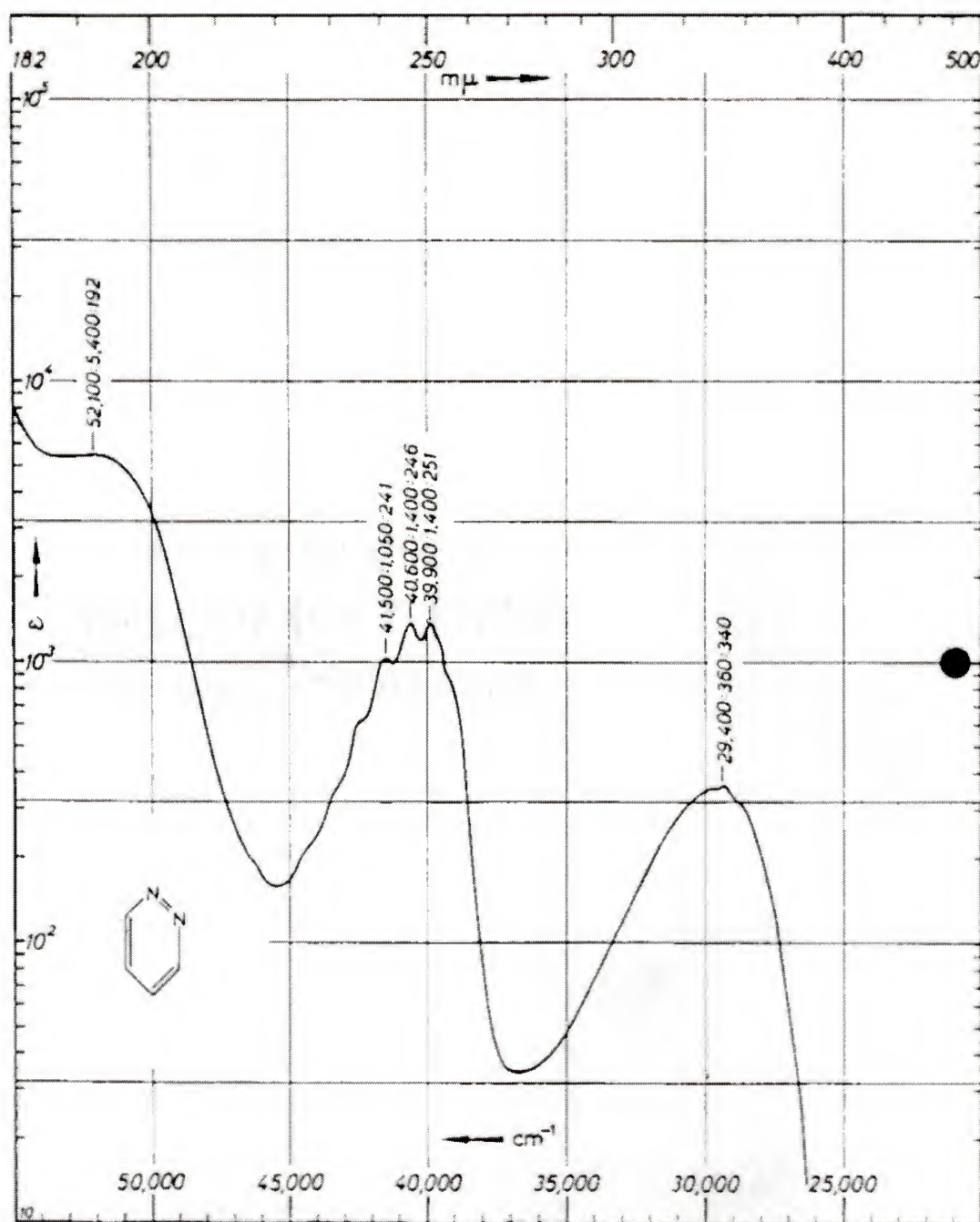


Figure 5.24: Experimental UV Spectra of Pyridazine in Non-polar Solvent

Reference 269

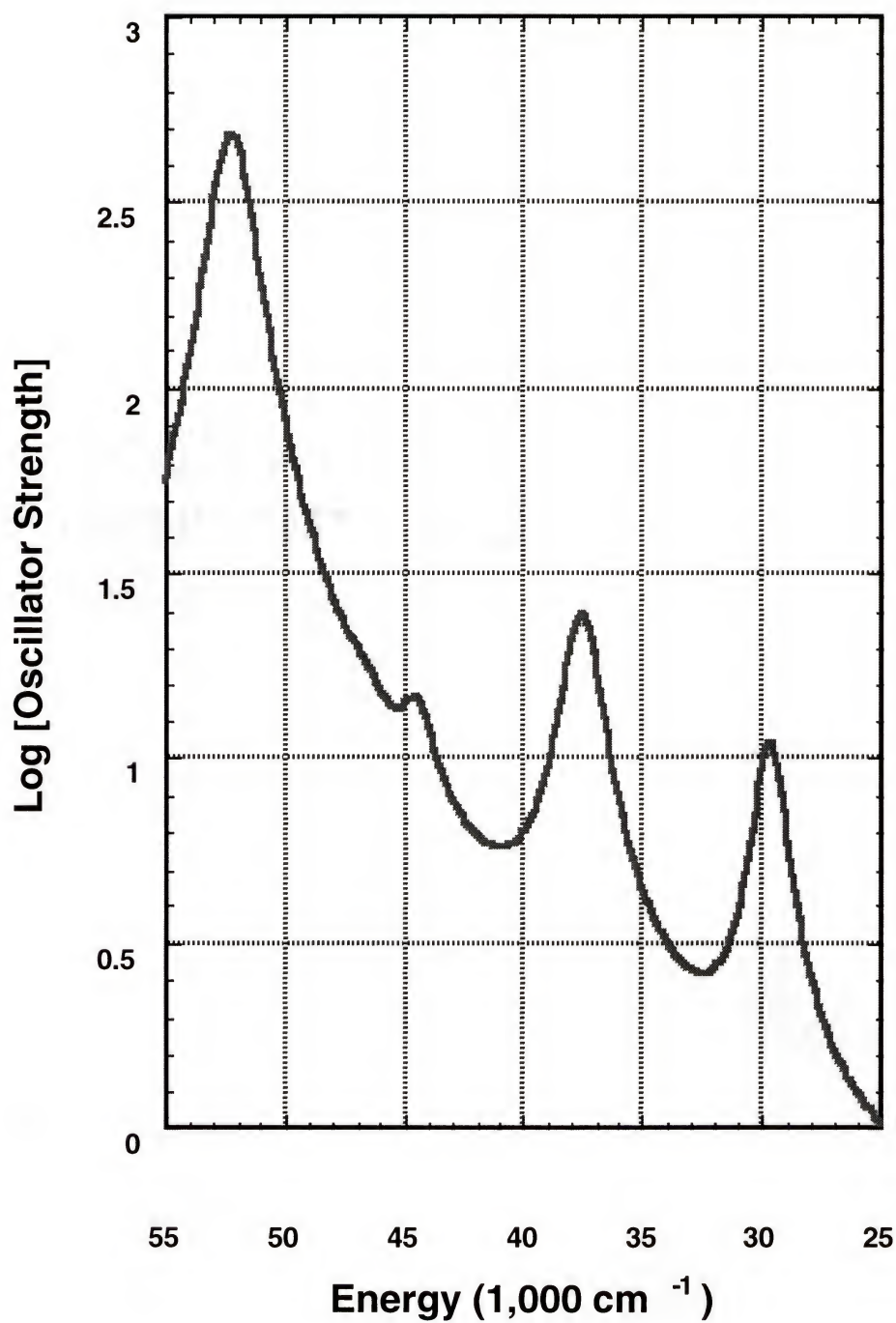


Figure 5.25: Calculated UV Spectra of Gas Phase Pyridazine Using the Band Width Theory

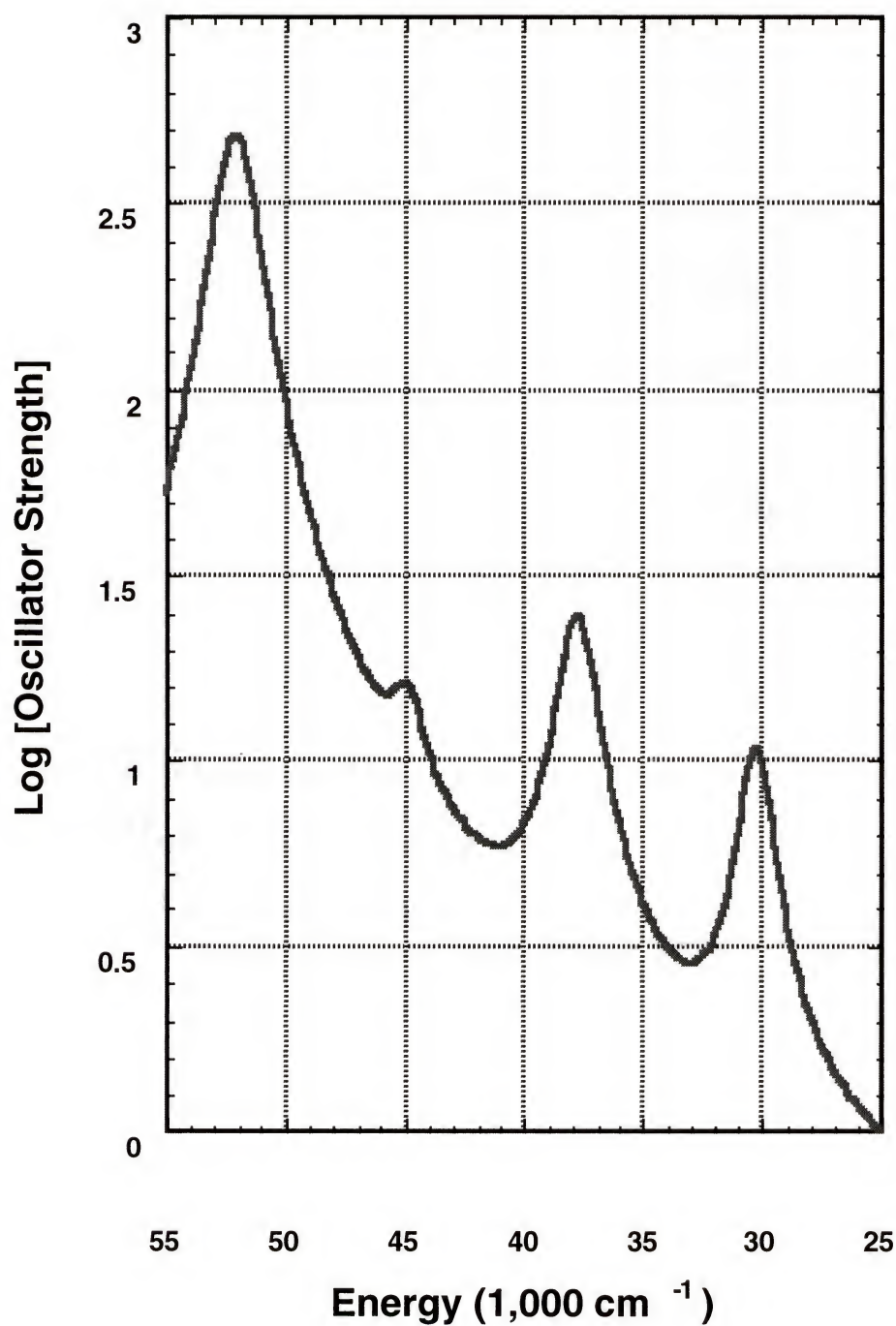


Figure 5.26: Calculated UV Spectra of Pyridazine in Non-polar Solvent Using the Band Width Theory

Spherical cavity SCRF calculation.

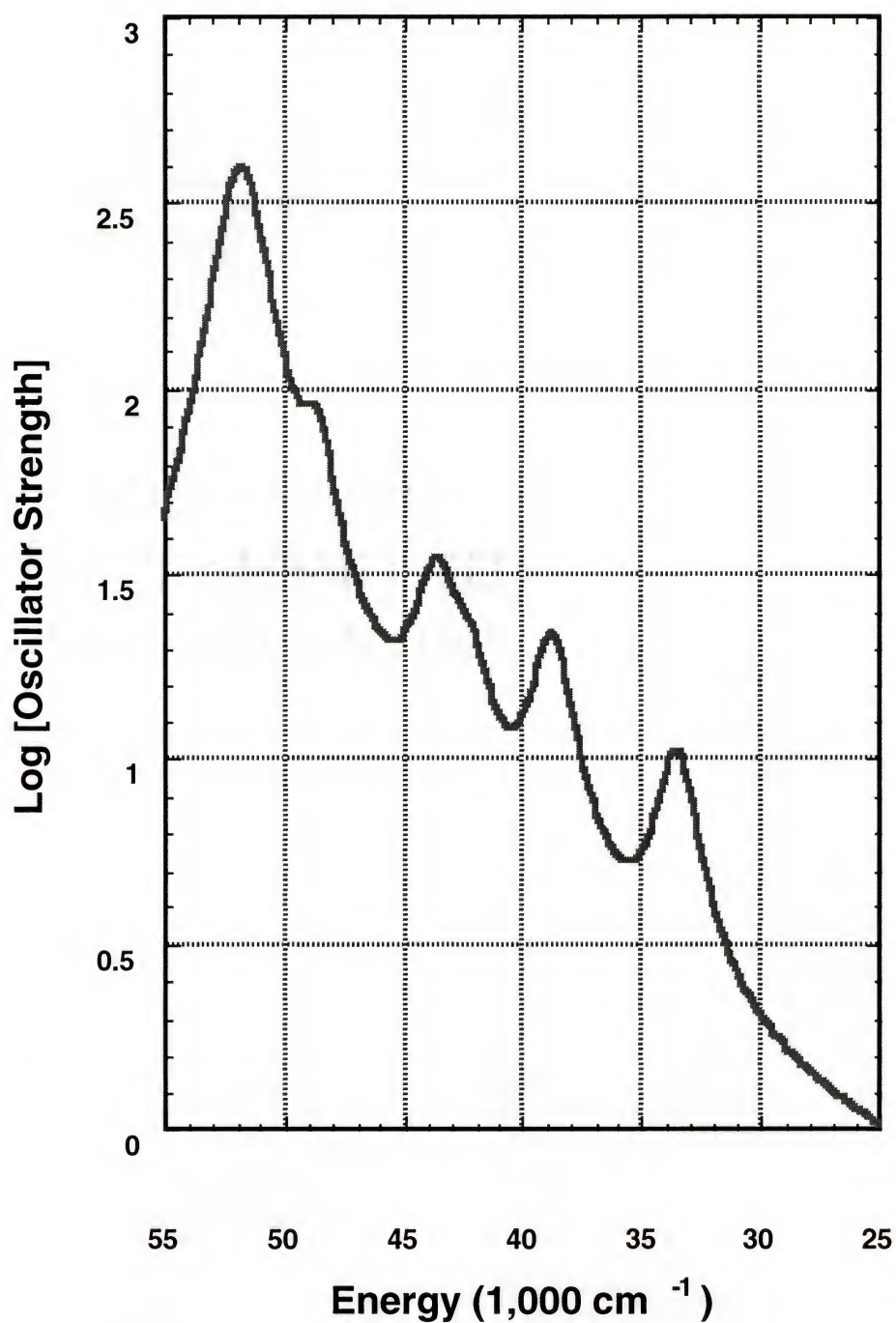


Figure 5.27: Calculated UV Spectra of Pyridazine Water Using the Band Width Theory

Spherical cavity SCRF calculation.

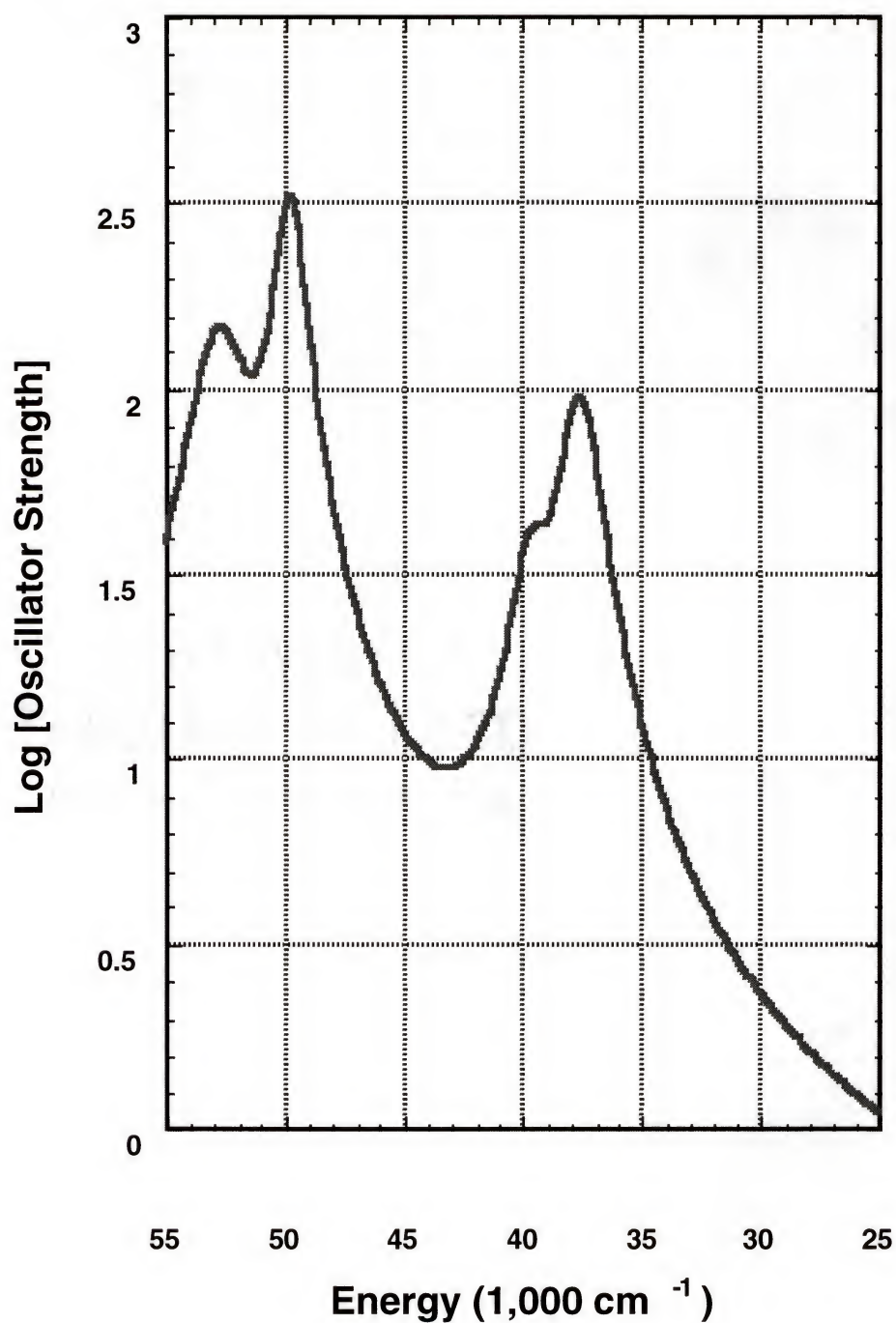


Figure 5.28: Calculated UV Spectra of Pyridazine Water Using the Band Width Theory 2

Spherical cavity SCRF calculation on supermolecule structure, nearest hydrogen bonding water included in the QM calculation.

Conclusion

A simple model for determining the band width for molecular electronic spectroscopy has been developed and examined.. The model has been shown to reproduce the appearance of the experimental spectra for small organic molecules. While this method does not reproduce the spectra as well as the stochastic method we describe, it does an adequate job and requires a fraction of the time. This model shows great promise for predicting the appearance of the spectra of large molecular systems in which stochastic calculations on such systems would simply be too expensive for the additional accuracy obtained.

This method has only two free parameters, that might be temperature and solvent dependent, although our preliminary studies suggest that this dependence is minor.

CHAPTER 6 SUMMARY AND CONCLUSIONS

Summary

The hybrid quantum mechanics-molecular mechanics (QM/MM) method, described within this thesis, was tested by comparing the results from high level *ab initio* calculations, experimental results, and other QM/MM methods for several hydrogen bonding dimers (Chapter 3). The results obtained from this implementation of a QM/MM method were more consistent than those published previously [120].

The first chemical problem described in this thesis examines the dynamics and photophysics of uracil and 1,3-dimethyluracil in protic and aprotic solutions (Chapter 4, Section 1). In the second problem we investigated the origin of the spectral sensitivity of ruthenium(II) pentaaminopyridine (Chapter 4, Section 2) to different solvents. In the third application I initiate an investigation of the binding energy of the hydrogen halide acids in aqueous solution (Chapter 4, Section 3). In the last example, the QM/MM method was used as a standard for verifying the results obtained from an empirical band width theory (Chapter 5).

The QM/MM method described herein has successfully demonstrated that this method can predict the solvent effects for a wide range of spectroscopic problems, some that the existing continuum models could not. The empirical band width method has provided a cost effective method for predicting the shape of a electronic transition from a single

energy calculation. This method is capable of modeling the solvent effects and the temperature dependence on UV spectra far more reliably, than are the simple reaction field methods.

Future Work

It would be interesting to consider application of the QM/MM method to large biological systems, such as the nitrogen transferase or the photochemical reaction center. These problems require the additional consideration of how to model a covalent bond when one atom is treated using quantum mechanics while the other atom in the bond is treated using classical molecular mechanics, the so called "linking atoms".

Since polarization of the environment is important, implementation of a polarizable force field into the hybrid QM/MM Hamiltonian is necessary in order to better represent the effects of the environment. Using a linear scaling method for the quantum mechanics partition would allow for a larger active, quantum mechanics region. These techniques could also allow the whole system to be treated using quantum mechanics methods periodically while searching phase space. This allows the classical point charges to change during the simulation, which would model some polarization in the classical region.

I would also like to explore stochastic methods for searching phase space. While the classical mechanics methods for searching phase space have been optimized for methods in which the energy calculation is similar in computer time as in the method for generating the new positions, this is not true when a quantum mechanics method is chosen for calculating the potential energy.

Implementing excited state gradients for the quantum mechanics region would also be of great interest. This would allow the QM/MM method not only to better simulate absorption spectra, but also allow the states of photochemical processes in solution. Additionally, this would allow modeling two electron absorption spectroscopy. This would also be useful in studying non-radiating mechanism for the excited states energy transfer.

APPENDIX A: ZINDO QM/MM User Manual

The QM/MM interface for the ZINDO program utilizes the same input file as the standard ZINDO program. This was accomplished by adding appropriate keywords to the existing input blocks, while placing the majority of the input specific to the QM/MM program in their own input blocks. This user manual is designed to function as a reference guide which should be used in conjunction with the ZINDO user manual. The schematic for the QM/MM program is given in Figure A1.1. The QM/MM program was designed to be somewhat modular, allowing easy implementation into other quantum mechanics programs, while utilizing existing ZINDO code.

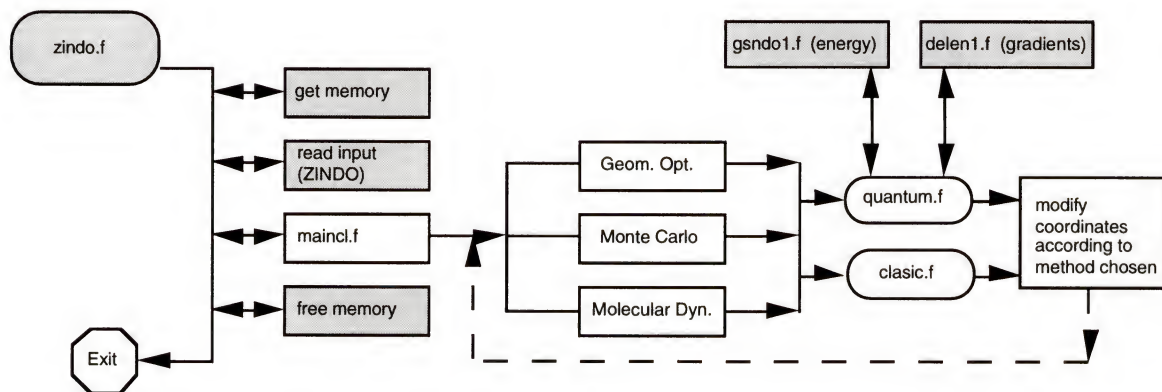


Figure A1.1: QM/MM Program Flow Chart

The shaded boxes represent the existing ZINDO code. This flow chart does not give a complete description of the QM/MM code or the ZINDO code, it does provide a general overview as to how the two programs are linked.

The current implementation of the QM/MM program can utilize most of the functionality and options available in ZINDO, the one exception to this is the PCM solvation model by Tomassi. This solvent model can not be used at the same time as the QM/MM program, since the previous manager of the ZINDO source code has these two programs using common arrays in the memory management.

\$CONTRL INPUT BLOCK

This is one of the current ZINDO control blocks, much of the information in this block is identical to the ZINDO program. The RUNTYP keyword is used to specify a QM/MM calculation, RUNTYP=QMMM. There are seven new keywords that can be used in this block if RUNTYP = QMMM, Table A1.1.

Table A1.1: \$CONTRL Keywords

Keyword	Description
QMMMTYP	Type of QM/MM Calculation
QMMMBND	Type of Boundary Conditions for Simulation
ENSEMBLE	Statistical Mechanics Ensemble
MMFF	Molecular Mechanics Force Field Method
LJHOLE	Method for Creating a Cavity in the Solvent
STEPSA	Atomic Step Size for Monte Carlo Simulation
STEPR	Molecular Step Size for Monte Carlo Simulation
MMCROT	Rotation Angle for the Monte Carlo Simulation

QMMMTYP

The QMMTYP keyword can have any of the following values:

G-SD	Geometry Optimization using a Steepest Descent algorithm; DEFAULT
G-BFGS	Geometry Optimization using the BFGS Hessian update algorithm.
G-DFP	Geometry Optimization using the DFP Hessian update algorithm.
G-GREEN	Geometry Optimization using the Green's Hessian update algorithm.
G-MUR	Geometry Optimization using the Mursar's Hessian update algorithm.
MD	Molecular Dynamics Simulation
MC	Metropolis Monte Carlo Simulation; Atomic Translations, Molecular Translations, and Molecular Rotations.
MC-MOLEC	Metropolis Monte Carlo Simulation; Atomic Translations only.
MC-RIGID	Metropolis Monte Carlo Simulation; Atomic Translations only
MC-RT	Metropolis Monte Carlo Simulation; Molecular Translations only
MC-ROT	Metropolis Monte Carlo Simulation; Molecular Rotations only
MC-TRAN	Metropolis Monte Carlo Simulation; Atomic and Molecular Translations.

QMMMBND

The QMMMBND keyword can have any of the following values:

NONE	No boundary conditions; DEFAULT
PBC	Periodic Boundary Conditions
MPBC	Minimum Periodic Boundary Conditions
SPHERE	Spherical Boundary Conditions (Not Functional)

ENSEMBLE

The ENSEMBLE keyword can have any of the following values:

NVT	Canonical Ensemble
NPT	Isothermal-Isobaric Ensemble
VT μ	Grand Canonical Ensemble (Not Functional)

MMFF

The MMFF keyword is used for specifying the functional form of the molecular mechanics potentials. The following values are acceptable:

CFF	Consistent Force Field
AMBER	AMBER force field
OPLS	OPLS united atom force field
MMX	Allinger's MMx family of force fields

LJHOLE

The LJHOLE keyword is used for creating a solvent hole in an equilibrated cluster of solvent molecules. There are two different methods implemented, The following are valid values for the LJHOLE;

- | | |
|---|--|
| 0 | No Cavity Creation; DEFAULT |
| 1 | Use a Lennard-Jones (12-6) Potential to create a spherical cavity in the solvent. |
| 2 | Create a hard sphere of the specified radius, this method is only useable with the Monte Carlo QMMM TYP's. |

STEPS

The STEPS keyword is used for specifying the step size that the Monte Carlo simulation uses for the atomic translations, the default value is 0.0025Å.

STEPR

The STEPR keyword is used for specifying the step size that the Monte Carlo simulation uses for the molecular translations, the default value is 0.1 Å.

MMCROT

The MMCROT keyword is used for specifying the step size that the Monte Carlo simulation uses for the molecular rotations, the default value is 5.0°.

Example Input Format:

\$CONRTL

SCFTYP	= ROHF	ENTTYP	= COORD
RUNTYP	= QMMM	SCFTOL	= 0.1D-10
QMMM TYP	= G-SD	UNITS	= ANG
APX	= INDO/1	INTTYP	= 0
QMMM BND	= NONE	MULT	= 1
NAT	= 4	NEL	= 12
PTCG	= 12	IDD2	= 0
NOP	= 2	NDT	= 1
MMFF	= CFF	ENSEMBLE	= NVT
FOP(1)	= 10.000000 2.000000		
ONAME	= NH3-4H2O		

\$END

\$DATAIN INPUT BLOCK

The \$DATAIN input block is identical to the ZINDO \$DATAIN input block. This block is used to read the coordinates for the quantum mechanics region only, do place point charges in this input block. Any of the input file formats that ZINDO supports will work, since the ZINDO program will read this information and then pass it along to the QM/MM program.

Example Input Format:

```
$DATAIN
X Y Z N END
...
$END
```

X,Y,Z are the coordinates while N is the atomic number (free format).


```

$DATAIN
-1.8200 0.6400 0.0000 8 END
-0.7700 0.0300 0.0000 6 END
0.1700 0.5700 0.0000 1 END
-0.7700 -1.0500 0.0000 1 END
$END

```

The END at the end of each line can be changed as indicated in the ZINDO input. Any valid form of \$DATAIN from the ZINDO manual can be used with the QM/MM program.

\$PTCHGI INPUT BLOCK

The \$PTCHGI input block is similar to the \$PTCHGI input block used by the standard ZINDO program. The only difference is that an additional integer value is required for each point charge. This additional value is the classical mechanics atom type, see section on molecular force fields.

Example Input Format:

```

$PTCHGI
TITLE LINE
"Screening Constant", "Bulk Dielectric"
X Y Z N Q AT
...
$END

```

X,Y,Z are the coordinates, N is the atomic number, Q is the effective charge, and AT is the molecular mechanics atom type (free format).

```

$PTCHGI
1 MM waters
0.00000 0.00000
-4.4500 -0.6000 0.0000 1 0.41 19
-4.4500 0.3600 0.0000 8 -0.82 18
-3.5500 0.6800 0.0000 1 0.41 19
$END

```

\$SCRFIN INPUT BLOCK

The \$SCRFIN input block is identical to the \$SCRFIN block used in ZINDO. This input block has an additional function though; it can be used to model the long range interactions. There are two different methods for doing this: first uses a spherical cavity with a radius equal to the cutoff radius; the second uses the multi-cavity for several fragments, again specifying the radius as the cutoff limit. The only restriction for this input block is that the PCM and COSMO solvent models do not function with the RUNTYP = QMMM, since they share some common arrays. Some of the former memory management practice might be changed to accommodate this. Both the keyword format and free format forms of the \$SCRFIN block can be used with the QM/MM program.

\$QCMDIN INPUT BLOCK

The \$QCMDIN input block is used for specifying information that pertains to molecular dynamics simulations. The following variables are used in this free format input block.

MSTEP	The number of time steps for this molecular dynamics simulation, maximum = 10,000.
OUTPUT	The frequency of how often the coordinates will be outputted to disk (MOVIE.xyz).
STEP	This is the step size used for the molecular dynamics simulations. 0.1D-12 - 0.1D-15 sec.
TEMP	This is the desired simulation temperature in Kelvin.

TAUTEMP	Friction Coefficient for the external heat bath, temperature scaling. This value usually varies from 0.1×10^{-12} - 0.1×10^{-15} , I usually use 0.7×10^{-12} .
NONBND	The molecular nonbonded cutoff radius (Å).
POTLIB	Name of the molecular mechanics force field file to be used for the simulation.
BOXA*	The X box length used for the periodic boundary conditions. THIS IS AN OPTIONAL VALUE, only required if QMMMBND \neq 0 in \$CONTRL.
BOXB*	The Y box length used for the periodic boundary conditions. THIS IS AN OPTIONAL VALUE, only required if QMMMBND \neq 0 in \$CONTRL.
BOXC*	The Z box length used for the periodic boundary conditions. THIS IS AN OPTIONAL VALUE, only required if QMMMBND \neq 0 in \$CONTRL.
SCALE6*	Optional Method for altering the LJ(6) interaction between the QM and MM regions. DEFAULT = 1.6. THIS IS AN OPTIONAL SWITCH.
SCALE12*	Optional Method for altering the LJ(12) interaction between the QM and MM regions. DEFAULT = 0.75. THIS IS AN OPTIONAL SWITCH.
IF LJHOLE = 1 in \$CONTROL	
ILJ	The atom type used for the Lennard-Jones(12) repulsion term.
SPHERE-RAD	The radius of the “soft” sphere
XORIGIN	The X origin for the sphere.
YORIGIN	The Y origin for the sphere.
ZORIGIN	The Z origin for the sphere.

Example Input Format:

There are two examples inputs for this block, the first has the required information only, while the second has some optional values set.


```
$QCMDIN
MSTEP OUTPUT
STEP TEMP TAUTEMP NONBND
POTLIB
$END
```

```
$QCMDIN
500 10
0.1E-15 200.0 0.7E-12 12.0
/ufl/qtp/mcz/progs/zindo/ForceFields/qcmd.lib
$END
```

```
$QCMDIN
MSTEP OUTPUT
STEP TEMP TAUTEMP NONBND
POTLIB
BOXA* BOXB* BOXC* SCALE6* SCALE12*
ILJ* SPHERE-RAD* XORGIN* YORGIN* ZORGIN*
$END
```

```
$QCMDIN
500 10
0.1E-15 200.0 0.7E-12 12.0
/ufl/qtp/mcz/progs/zindo/ForceFields/qcmd.lib
20.0* 20.0* 20.0* 1.0* 1.0*
1* 5.0* 1.0* 2.0* 1.0*
$END
```

* is the optional input variables.

\$QCGIN INPUT BLOCK

The \$QCGIN input block is used for specifying information that pertains to geometry optimization procedures. The following variables are used in this free format input block.

STEP	The size of the steepest descent step (Å), needed for the initial step used in Hessian update methods.
ETOL	When the difference in energy (kJ/mol) is less than this number the optimization is consider complete.

GTOL	This is used for testing the gradient, when the gradient is less than this number the optimization is complete.
CUTOFF	The molecular nonbonding cutoff radius.
SCALE6*	Optional Method for altering the LJ(6) interaction between the QM and MM regions. DEFAULT = 1.6
SCALE12*	Optional Method for altering the LJ(12) interaction between the QM and MM regions. DEFAULT = 0.75
MSTEP	Maximum number of geometry optimization steps to perform.
POTLIB	Name of the molecular mechanics force field file to be used for the simulation.
BOXA *	The X box length used for the periodic boundary conditions.
BOXB*	The Y box length used for the periodic boundary conditions.
BOXC*	The Z box length used for the periodic boundary conditions.
IF LJHOLE = 1 in \$CONTROL	
ILJ*	The atom type used for the Lennard-Jones(12) repulsion term.
SPHERE-RAD*	The radius of the "soft" sphere
XORIGIN*	The X origin for the sphere.
YORIGIN*	The Y origin for the sphere.
ZORIGIN*	The Z origin for the sphere.

Example Input Format:

There are two examples inputs for this block, the first has the required information only, while the second has some optional values set.

```
$QCGIN
STEP ETOL GTOL CUTOFF
MSTEP
POTLIB
$END
```

```
$QCGIN
0.1 0.001 0.0001
50
/ufl/qtp/mcz/progs/zindo/ForceFields/qcmd.lib
$END
```

```
$QCGIN
STEP ETOL GTOL CUTOFF SCALE6* SCALE12*
MSTEP
POTLIB
BOXA* BOXB* BOXC*
ILJ* SPHERE-RAD* XORGIN* YORGIN* ZORGIN*
$END
```

```
$QCGIN
0.025 0.0001 0.00001 1.5*
500
/ufl/qtp/mcz/progs/zindo/ForceFields/qcmd.lib
200.0* 200.0* 10.0*
1* 1.0*
```

* is the optional input variables.

\$QCMCIN INPUT BLOCK

The \$QCMCIN input block is used for specifying information that pertains to Monte Carlo simulations. The following variables are used in this free format input block.

TEMP	This is the desired simulation temperature in Kelvin.
NONBND	The molecular nonbonded cutoff radius (Å).
SCALE6*	Optional Method for altering the LJ(6) interaction between the QM and MM regions. DEFAULT = 1.6

SCALE12*	Optional Method for altering the LJ(12) interaction between the QM and MM regions. DEFAULT = 0.75
ICONF	The number of accepted configuration needed.
OUTPUT	The frequency of how often the coordinates will be outputted to disk (MOVIE.xyz).
POTLIB	Name of the molecular mechanics force field file to be used for the simulation.
BOXA*	The X box length used for the periodic boundary conditions.
BOXB*	The Y box length used for the periodic boundary conditions.
BOXC*	The Z box length used for the periodic boundary conditions.
IF LJHOLE = 1 in \$CONTROL	
ILJ*	The atom type used for the Lennard-Jones(12) repulsion term.
SPHERE-RAD*	The radius of the “soft” sphere
XORGIN*	The X origin for the sphere.
YORGIN*	The Y origin for the sphere.
ZORGIN*	The Z origin for the sphere.
IF LJHOLE = 2 in \$CONTROL	
SPHERE-RAD*	The radius of the “hard” sphere
XORGIN*	The X origin for the sphere.
YORGIN*	The Y origin for the sphere.
ZORGIN*	The Z origin for the sphere.

Example Input Format:

There are two examples inputs for this block, the first has the required information only, while the second has some optional values set.

```
$QCMCIN
TEMP NONBND
ICONF OUTPUT
POTLIB
$END
```

```
$QCMCIN
210.5 12.5
200000
/uf1/qtp/mcz/progs/zindo/ForceFields/qcmd.lib
$END
```

```
$QCMCIN
TEMP NONBND SCALE6* SCALE12*
ICONF OUTPUT
POTLIB
BOXA* BOXB* BOXC*
ILJ* SPHERE-RAD* XORGIN* YORGIN* ZORGIN*
      or
SPHERE-RAD* XORGIN* YORGIN*
$END
```

```
$QCMCIN
210.5 12.5 2.0* 3.5*
200
/uf1/qtp/mcz/progs/zindo/ForceFields/qcmd.lib
20.0 20.0 20.0
$END
```

* is the optional input variables.

\$CLASIC INPUT BLOCK

The \$CLASIC input block is used for specifying which atoms belong to a molecule, which atoms are in the quantum or classical region, bond list for the classical mechanics, if the atom is fixed in phase space (geom opt and Monte Carlo only). This input block is free format, although it is not keyword driven. The following variables are needed:

RESTART	This character variable is used to specify if this is a new simulation of a continuation of an previous simulation. If "OLD" is specified CMRSTR, QMRSTR, and FORCE files are needed in the directory that the input file resides in. The FORCE file is not required for Monte Carlo simulations. These files are always printed out at the end of a simulation.
NATOM	This is the total number of atoms in a simulation, (QM + MM)
NMOL	This is the total number of molecules in a simulation, (QM + MM)
NMQM	This is the number of quantum mechanics molecules in the simulation. When using the SCRF multi-cavity method this should equal the number of fragments.
NMCL	This is the number of classical mechanics molecules in the simulation.
ATOM#	This is the atom number for each atom, this is important for internal accounting.
MOL#	This is the molecule that the specified atom belong to, this is important for internal accounting.
QM/MM	<p>This variable specifies if the atom is to be treated as quantum mechanics or molecular mechanics. The following value are valid:</p> <ul style="list-style-type: none"> 1 = Quantum Mechanics 2 = Classical Mechanics 3 = Constrained Quantum Mechanics 4 = Constrained Classical Mechanics 5 = Quantum Mechanics Link Atom 6 = Quantum Mechanics Link Bond <p>the values for 5 and 6 are not currently functional as of Spring 1998.</p>
POT	<p>This allows you to explicitly alter the force field parameters without having to modify the force field library.</p> <ul style="list-style-type: none"> 0 = Read parameters from Force Field file. 1 = Read parameter from \$PARAM Input Block <p>Not completely implemented.</p>

BOND

This is a set of 6 integers that specify which atoms are bonded to this atom. Only specify atom#'s that are higher the current atom number.

This field is used to specify the classical mechanics atom type for quantum mechanics atoms.

Example Input Format:

```
$CLASIC
RESTART
NATOM NMOL NMQM NMCM
ATOM# MOL# QM/MM POT (BOND) BOND BOND BOND BOND BOND
ATOM# MOL# QM/MM POT (BOND) BOND BOND BOND BOND BOND
...
$END
```

The value for (BOND) depends if the atom is QM or MM.

```
$CLASIC
NEW
6 4 1 1
1 1 1 0 19 (QM)
2 1 3 0 18 (QM)
3 1 1 0 19 (QM)
4 2 2 0 5 0 0 0 0 0 (MM)
5 2 2 0 6 0 0 0 0 0 (MM)
6 2 2 0 0 0 0 0 0 0 (MM)
$END
```

Molecular Mechanics Force Fields

The default force field file for the ZINDO QM/MM program is located in `~mczprogs/zindo/ForceFields/qcmd.lib`. The QM/MM program can be interfaced to any molecular mechanics force field, although some coding may have to be done to accommodate nonstandard molecular mechanics functional. The parameters for Atom Type 18 and higher should be used with caution, since they are not part of the standard CFF force field parameters. Also you need to determine the charges for the systems being investigated, the force field does not include a standard set of charges. You should make a copy of this force field and modify your own copy, and only add parameters that have been tested. The MMFF keyword in the \$CONTRL block determines the format needed for reading in the force field file, see *readparam.f*, in the *clasic* subdirectory of the ZINDO program. It is easier to add new force field types than modifying the force fields into another format, see *getvar.f* in the *io* subdirectory of the ZINDO program for modifying the MMFF keyword.

Table A1.2: Default Molecular Mechanics Force Field (CFF)

Atom Type	Description	Atom Type	Description
1	Carbonyl Oxygen	20	Sodium Ion
2	Hydroxyl Oxygen	21	Ruthenium
3	Carboxyl Oxygen	22	Iron +2
4	Nitrogen	23	Iron +3
5	1/3 Charged Nitrogen	24	Iron
6	NH3 Terminal Group	25	Fluorine
7	Aliphatic CH Carbon	26	Chlorine
8	Aliphatic CH2 Carbon	27	Bromine
9	Aliphatic CH3 Carbon	28	Iodine
10	Bare Carbon	29	Magnesium
11	Aromatic Carbon	30	Copper
12	Sulfur	31	TIPS Oxygen
13	Hydrogen	32	TIPS Hydrogen
14	Nonpolar Hydrogen	33	Nitrogen in Ammonia
15	Aromatic Nitrogen	34	Hydrogen in Ammonia
16	Double Bonded Nitrogen	35	Hydrogen in Chloroform
17	Calcium Ion	36	Carbon in Chloroform
18	SPC Oxygen in Water (Q = -0.82)	37	Chlorine in Chloroform
19	SPC Hydrogen in Water (Q = 0.41)		

Computer Code Description

angle.f	This subroutine calculates the bond angle energies and forces for the classical region.
bond.f	This subroutine calculates the bond distance energies and forces for the classical region.
boundary.f	This subroutine is used for the periodic boundary conditions.
centmass.f	This subroutine calculates the center of mass for each molecule
classical.f	This is the main control subroutine for calculating the classical mechanics energies and forces.
cmmoment.f	This subroutine calculates the higher electrical moments for the classical mechanics region.
cmpbc.f	This subroutine calculates the non-bonding energies and forces for the classical region.
cmpbc2.f	This subroutine is used for Monte Carlo simulations, it only calculates the Lennard-Jones (12-6) energies.
coormove.f	This subroutine is the corrector step in molecular dynamics simulations.
dcmqmcoul.f	This subroutine calculates the derivatives for the QM/MM Coulombic interactions for the MM region.
dcmqmvdw.f	This subroutine calculates the derivatives for the QM/MM Lennard-Jones (12-6) interactions for the MM region.
dihed.f	This subroutine calculates dihedral angles
dqmcmvdw.f	This subroutine calculate the derivatives for the QM/MM Lennard-Jones (12-6) interactions for the QM region.
elj.f	This subroutine is 2 functions, one calculate the energy for Lennard-Jones (12-6) potential whereas the other calculates the derivative.
geomhess.f	This subroutine is used for geometry optimization methods that require Hessian's

geomopt.f	This subroutine controls the geometry optimization calculations for the QM/MM simulations.
geomsd.f	This subroutine is used or calculating the GNORM and steepest descent geometry optimization.
maincl.f	This is the main program for the QM/MM program, it allocates all of the memory necessary for the QM/MM module.
mcadjust.f	This subroutine adjusts the step sizes of the Monte Carlo simulation to maintain the optimal acceptance ratio.
mcenergy.f	This subroutine calculates the classical mechanics energy for the Monte Carlo simulations.
mcmove.f	This subroutine changes the working coordinates for a Monte Carlo simulation after it has been accepted.
mctest.f	This subroutine test to see if the new configurations is accepted or not.
mcupdate.f	This subroutine updates all the information for the specified statistical mechanics ensemble.
mdpredict.f	This subroutine is used for predicting the new coordinates to used for the Monte Carlo simulation.
mkhole.f	This subroutine is used for the soft sphere cavity model.
monte.f	This subroutine controls the Monte Carlo simulations.
moveatom.f	This subroutine is used for Monte Carlo simulations for atomic translations.
movemol.f	This subroutine is used for Monte Carlo simulations for molecular translations.
qcmd.f	This subroutine controls the molecular dynamics simulations.
qcmoment.f	This subroutine is called on the last time step of molecular dynamics simulations to provide statistic for the simulation.
qcmcmoul.f	This subroutine is used to calculate the QM/MM Coulombic interaction energy for the QM region.
qcmcmvdw.f	This subroutine is used to calculate the QM/MM

Lennard-Jones (12-6) interaction energy for the QM region

quantum.f	This is the main subroutine for calculating the quantum mechanics energies and forces.
readin.f	This subroutine is used to read in and set up the input file for the QM/MM simulation.
readparam.f	This subroutine is used for reading in the force field library file and setting up the classical mechanics equations.
rotmol.f	This subroutine is used for Monte Carlo simulations for molecular translations.
rstrt.f	This subroutine produces a restart file.
scalet.f	This subroutine is used for the molecular dynamics simulations to scale the temperature.
speed.f	This subroutine is used to calculate the velocities for a MD simulation.
spherebc.f	This subroutine is used for the spherical boundary conditions (stochastic).
stapos.f	This subroutine is used for starting new molecular dynamics simulations.
tempture.f	This subroutine is used to calculate the temperature for molecular dynamics simulations.
testhole.f	This subroutine is used for the hard sphere cavity option.
torsion.f	This subroutine is the original subroutine for calculating the torsion energy and forces for the classical region.
torsion2.f	This subroutine is used for the AMBER force field torsion energy and forces, different functional form than torsion.f

APPENDIX B: STOCHASTIC SIMULATION ANALYZER MANUAL

The QM/MM analyzer program is used for analyzing the data obtained from the simulation. This program is designed to use the MOVIE.xyz file, created automatically by the ZINDO QM/MM program, as the input file. The QM/MM analyzer program is designed as an interactive program, and hence the MOVIE.xyz file can be renamed. There is on on-line HELP available for this program, although it is rather crude. The following is designed to be a reference manual. Figure A2.1 is the flow chart for the QM/MM analyzer program. Since the program is interactive and includes on-line help, a description of how to run the program is not included herein. A brief description of each subroutine is provided although.

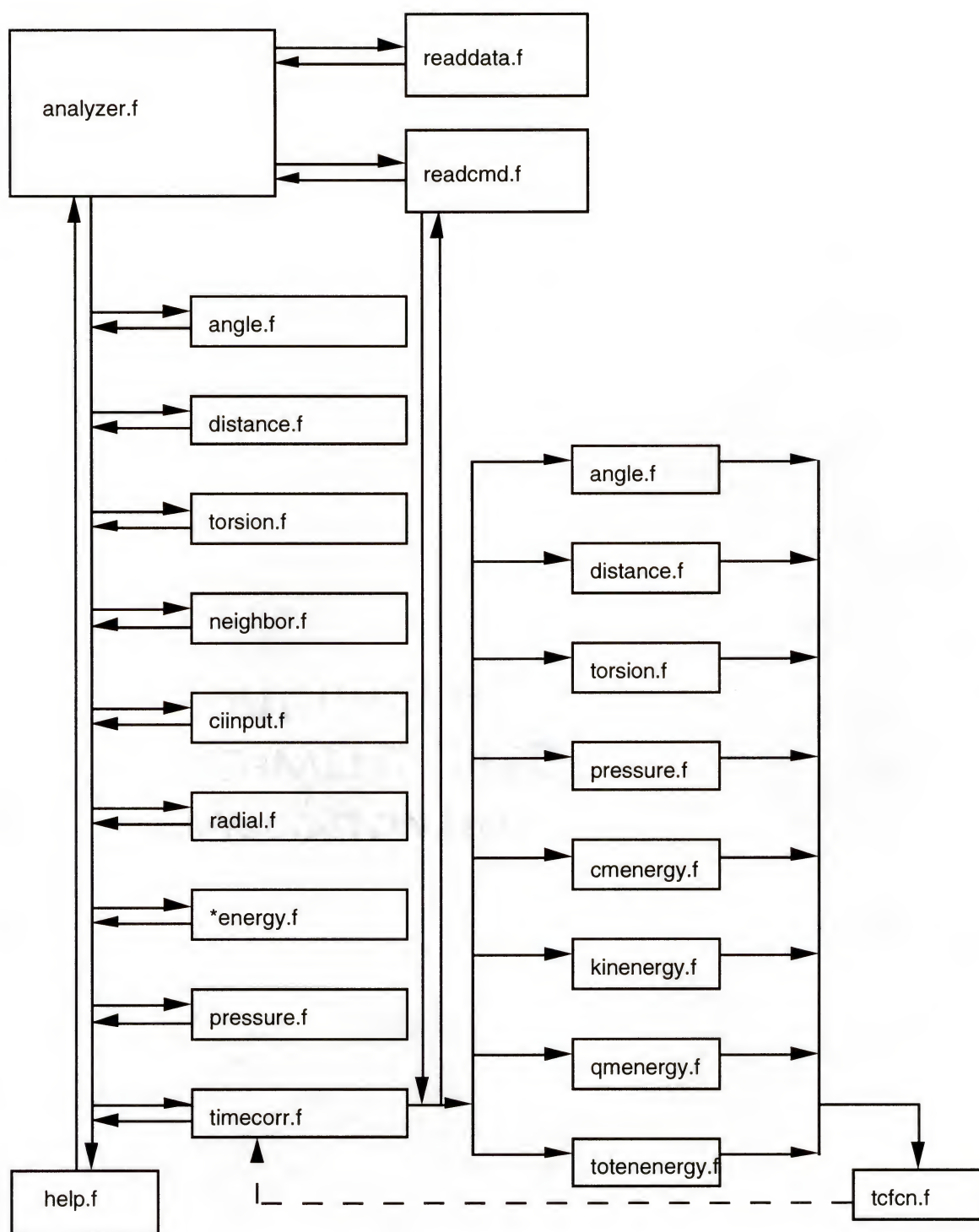


Figure A2.1: Flow Chart for QM/MM Analyzer Program.

Computer Code Description

analyzer.f	This is the main program, this program is required to use FORTRAN 90 (F90), since it uses F90 memory allocations calls.
angle.f	This subroutine is used for measuring a specified angle for the indicated time period.
ciinput.f	This subroutine is used for creating ZINDO input files, primarily for CIS calculations.
cmenergy.f	This subroutine is used for calculating the average energy of the classical partition for the specified time period.
density.f	This subroutine is used for calculating the density.
distance.f	This subroutine is used for measuring the specified distance for the indicated time period.
help.f	This subroutine provides on-line help, list the appropriate keywords and provides some information as to the functionality.
kinenergy.f	This subroutine is used for calculating the average kinetic energy for the specified time period.
neighbor.f	This subroutine is used for listing the atoms that are within the specified radius from the desired atom.
pressure.f	This subroutine is used for calculating the pressure for the simulation for the specified time period.
qmenergy.f	This subroutine is used for calculating the average energy for the quantum mechanics partition for the specified time period.
radial.f	This subroutine is used for calculating radial distributions functions.
readcmd.f	This subroutine is used for reading the commands entered. This subroutine is also responsible for transforming upper case letters to lower case.

readdata.f	This subroutine is used for processing the input file (MOVIE.xyz) and storing the information in memory.
tcfcn.f	This subroutine is used for calculating the time correlation function.
timecorr.f	This subroutine is used for determining the type of time correlation function to calculate.
torsion.f	This subroutine is used for calculating the average torsion angle for the specified time interval.
totenergy.f	This subroutine is used for calculating the average total energy of the simulated system for the specified time period.

REFERENCES

- 1 G.S. Beddard; *Philos. T. Roy. Soc. A*, **102**, 421 (1998)
- 2 R.K. Clayton; *Photosynthesis: Physical mechanisms and chemical patterns, volume 4 of IUPAB Biophysics Series*, chapter 1,9, pages 1-19, 191-226, Cambridge University Press, (1980).
- 3 M. Bixon, J. Jortner, M. Plato, and M.E. Michel-Beyerle; *The Photosynthetic Bacterial Reaction Center Structure and Dynamics*, pages 399-419, NATO ASI Series, Plenum, (1988).
- 4 R.A. Marcus; *The Photosynthetic Bacterial Reaction Center Structure and Dynamics*, pages 389-398, NATO ASI Series, Plenum, (1988).
- 5 M.A. Thompson and M.C. Zerner; *J. Am. Chem. Soc.*, **113**, 8210 (1991).
- 6 R.B. Silverman; *The Organic Chemistry of Drug Design and Drug Action*, Academic Press, San Diego, (1992).
- 7 L. Stryer; *Biochemistry*, W.H. Freeman and Company, New York, (1988).
- 8 A. Streitwieser, C.H. Heathcock, and E. M. Kosower; *Introduction to Organic Chemistry*, Macmillan Publishing Company, New York, (1992).
- 9 R.H. Petrucci and W.S. Harwood; *General Chemistry: Principles and Modern Applications*, Prentice Hall, New Jersey, (1997).
- 10 *Chemical and Engineering News*; annual report on top manufactured chemicals, (1997).
- 11 D.B. Boyd; *Reviews in Computational Chemistry*, Vol. 5, editors K.B. Lipkowitz and D.B Boyd, VCH, New York, (1995).
- 12 G.M. Downs and P. Willett; *Reviews in Computational Chemistry*, Vol. 6, editors K.B. Lipkowitz and D.B Boyd, VCH, New York, (1996).

- 13 A. Szabo and N.S. Ostlund; *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill Publishing Company, New York, (1989).
- 14 P.W. Atkins; *Physical Chemistry*, W.H. Freeman and Company, San Francisco, (1978).
- 15 R.J. Bartlett, N.Y. Öhrn, N. Oliphant, N. Richards, J.F. Stanton, J.D. Watts, and M.C. Zerner; *Lecture Notes from the Florida School on Applied molecular orbital Theory*, University of Florida, (1992).
- 16 J.P. Bowen and N.L. Allinger; *Reviews in Computational Chemistry, Vol 2*, editors K.B. Lipkowitz and D.B. Boyd, VCH, New York, 1991.
- 17 K.B. Lipkowitz, G. Pearl, B. Conor, and M.A. Peterson; *J. Am. Chem. Soc.*, **119**, 600 (1997).
- 18 J.M. Briggs, T.B. Nguyen, W.L. Jorgensen; *J. Phys. Chem.*, **95**, 3315 (1991).
- 19 C.E. Dykstra; *Quantum Chemistry and Molecular Spectroscopy*, Prentic Hall, New Jersey, (1992).
- 20 E. Arunan, C.E. Dykstra, and H.S. Gutowsky; *J. Chem. Phys.*, **105**, 8495 (1996).
- 21 S. Jun-Qiang and R.J. Bartlett; *J. Chem. Phys.*, **108**, 301 (1998).
- 22 M.L. Leininger and H.F. Schaefer III; *J. Chem. Phys.*, **107**, 9059 (1997).
- 23 H.F. Scafer III, Ed.; *Modern Theoretical Chemistry, Vol. 3 and 4*, Plenum Press, New York, 1977.
- 24 W.J. Hehre, L. Random, P.v.R. Schleyer, and J.A. Pople; *Ab Initio Molecular Orbital Theory*, John Wiley & Sons, New York, 1986.
- 25 A.E.S. Green and M. Zanardi; *Int. J. Quantum Chem.*, **66**, 219 (1998).
- 26 C. Reichardt; *Solvent s and Solvent Effects in Organic Chemistry*, VCH publishers, New York, 1990.
- 27 W.P. Hayes and C.J. Timmons; *Spectrochim. Acta*, **21**, 529 (1965).
- 28 M.B. Ledger and P. Suppan; *Spectrochim. Acta, Part A*, **23**, 641 (1967).

- 29 M.S. Walker, T.W. Bendar, and R. Lumry; *J. Chem. Phys.*, **47**, 1020 (1967).
- 30 M.J. Field; *Computer Simulations of Biomolecular Systems: Theoretical and Experimental Applications, Vol 2*, editors, W.F. van Gunsteren, P.K. Weiner, and A.J. Wilkinson, ESCOM, Leiden, 1993.
- 31 J. Chandrasekhar, S.F. Smith, and W.L. Jorgensen; *J. Am. Chem. Soc.*, **107**, 154 (1985).
- 32 T.P. Lybrand, I. Ghosh, and J.A. McCammon; *J. Am. Chem. Soc.*, **107**, 7793 (1985).
- 33 J.A. McCammon; *Science*, **238**, 486 (1987).
- 34 S.E. DeBolt and P.A. Kollman; *J. Am. Chem. Soc.*, **113**, 4491 (1991).
- 35 W.L. Jorgensen; *J. Am. Chem. Soc.*, **103**, 335 (1981).
- 36 W.L. Jorgensen; *J. Chem. Phys.*, **77**, 4156 (1982).
- 37 J.R. Reimers, R.O. Watts, and M.L. Klein; *Chem. Phys.*, **65**, 95 (1982).
- 38 R.W. Impey and M.L. Klein; *Chem. Phys. Lett.*, **104**, 579 (1984).
- 39 W.L. Jorgensen; *J. Chem. Phys.*, **77**, 4156 (1982).
- 40 K. Coutinho, S. Canuto and M.C. Zerner; *Intern. J. Quantum Chem.*, **65**, 885 (1997).
- 41 A.A. Rashin and B. Honig; *J. Phys. Chem.*, **89**, 5588 (1985).
- 42 A.A. Rashin and K. Namboodiri; *J. Phys. Chem.*, **91**, 6003 (1987).
- 43 M.M. Karelson, T.Tamm, A.R. Katriyzky, S.J. Cato, and M.C. Zerner; *Tetrahedron Comput. Method.*, **2**, 295 (1989).
- 44 A. Klamt and G. Schüürmann; *J. Chem. Soc., Prekins Trans.*, **2**, 799 (1993).
- 45 C.J. Cramer and D.G. Truhlar; *Science*, **256**, 213 (1992).
- 46 B. Mennucci, M. Cossi, and J. Tomasi; *J. Chem. Phys.*, **102**, 6837 (1995).
- 47 O. Tapia and G. Johannin; *J. Chem. Phys.*, **75**, 3624 (1981).

- 48 J.H. McCreery , R.E. Christoffersen, and G.G. Hall; *J. Am. Chem. Soc.*, **98**, 7198 (1976).
- 49 B.T. Thole and P. Th. van Duijnen; *Theor. Chim Acta*, **55**, 307 (1980).
- 50 W.C. Still, A. Tempczyk, R.C. Hawley, and T. Hendrickson; *J. Am. Chem. Soc.*, **112**, 6127 (1990).
- 51 H. Nagae; *J. Chem. Phys.*, **106**, 5159 (1997).
- 52 R. Marcus; *J. Chem. Phys.*, **24**, 966 (1956).
- 53 M.M. Karelson and M.C. Zerner; *J. Phys. Chem.*, **96**, 6949 (1992).
- 54 A.H. de Vries, P. Th. van Duijnen, and A.H. Juffer; *Int. J. Quantum Chem., Quantum Chem. Symp.*, **27**, 451 (1993).
- 55 C.J. Cramer and D.G. Trular; *J. Am. Chem. Soc.*, **113**, 8305 (1991).
- 56 C.J. Cramer and D.G. Truhlar; AMSOL (Program 606
Quatum Chemistry Program Exchange, Indiana University,
Bloomington, IN).
- 57 C.J. Cramer and D.G. Truhlar; *J. Comput. Chem.*, **13**, 1089 (1992).
- 58 M.C. Zerner; *Navel Research Rieviews*, Four/1993 - One/1994
- 59 A. Warshel and M. Levitt; *J. Mol. Biol.*, **103**, 227 (1976).
- 60 J. Gao; *Reviews In Compuatational Chemistry, Vol 7,*
Chapter 3, editors K.B. Lipkowitz and D.B. Boyd, VCH
Publishers Inc., New York, 1996.
- 61 J. Åqvist and A. Warshel; *Chem. Rev.*, **93**, 2523 (1993).
- 62 K.P. Eurenus, D.C. Chatfield, and B.R. Brooks; *Int. J. Quantum Chem.*, **60**, 1189 (1996).
- 63 D. Jacquemin, J.A. Morales, E. Deumens, and Y. Öhrn; *J. Chem. Phys.*, **107**, 6146 (1997).
- 64 A. Brünger, M. Karplus, and R. Huber; *Biochemistry* **26**, 5153 (1987).

- 65 ZINDO-97, a semi-empirical program package. M.C. Zerner, C. Warren, J.E. Ridley, A.D. Bacon, R.F. Kirchner, W.D. Edwards, J. McKelvey, J.C. Culberson, J.D. Head, B. Weimer, A.D. Cameron, M. Kotzian, N. Roesch, B. Parkinson, J. Yu, R. Pitzer, M. Cory, T. Tamm, M. Karelson, and G. Pearl; University of Florida, Gainesville, FL 32611-8435.
- 66 D.H. Andrews; *Phys. Rev.* **36**, 544 (1930).
- 67 N.L. Allinger; *Adv. Phys. Org. Chem.* **13**, 1 (1976).
- 68 S.R. Niketic and K. Rasmussen; *The Consistent Force Field*, Springer, New York, 1977.
- 69 N.L. Allinger, Y.H. Yuh, and J.-H. Lii; *J. Am. Chem. Soc.* **111**, 8551 (1989).
- 70 A. Warshel and S. Lifson; *J. Chem. Phys.* **49**, 5116 (1968).
- 71 A. Warshel and S. Lifson; *J. Chem. Phys.* **53**, 582 (1970).
- 72 S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Proteta, and P. Weiner; *J. Am. Chem. Soc.* **107**, 3902 (1985).
- 73 B.R. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, and M. Karplus; *J. Comp. Chem.* **4**, 187 (1983).
- 74 Mayo, Olafson, and Goddard; *J. Phys. Chem.* **94**, 8897 (1990).
- 75 H.J.C. Berendsen, J.P.M Postma, W.F. von Gunstarin, and J. Hermans; *Intermolecular Forces*, edited by B. Pullman, Reidel, Dordrecht, Holland, 1981.
- 76 J.E. Lennard-Jones; *Proc. Phys. Soc. (London), Ser. A* **43**, 461 (1931).
- 77 E. Schrödinger; *Ann. Phys.* **79**, 361 (1926).
- 78 E. Schrödinger; *Ann. Phys.* **81**, 109 (1926).
- 79 W. Heisenberg; *Z. Phys.* **33**, 879 (1925).
- 80 W. Heitler and F. London; *Z. Physik* **44**, 455 (1927).
- 81 M. Born and R. Oppenheimer; *Ann. Phys.* **84**, 457 (1927).
- 82 D. Hartree; *Proc. Cambridge. Phil. Soc.* **24**, 89,111,426 (1928).
- 83 F. Hund; *Z. Physik* **61**, 126 (1928).

- 84 R. Mulliken; *Phys. Rev* **32**, 186,388,761 (1928).
- 85 R. Mulliken; *Phys. Rev* **33**, 730 (1929).
- 86 V. Fock; *Z. Physik* **61**, 126 (1930).
- 87 R. Mulliken; *Phys. Rev* **40**, 55 (1932).
- 88 R. Mulliken; *Phys. Rev* **41**, 149,751 (1932).
- 89 T. Koopmans; *Physica* **1**, 104 (1933).
- 89 R.S. Mulliken; *J. Chem. Phys.* **2**, 782 (1934).
- 90 D. Hartree, W. Hartree, and B. Swirles; *Phil. Trans. Roy. Soc. (London)* **A238**, 229 (1939).
- 91 C.C.J. Roothaan; *Rev. Mod. Phys.* **23**, 69 (1951).
- 92 R.S. Mulliken; *J. Chem. Phys.* **23**, 1833 (1955).
- 93 C.C.J. Roothaan; *Rev. Mod. Phys.* **32**, 179 (1960).
- 94 W. Kohn and L.J. Sham; *Phys. Rev.* **140**, 1133 (1965).
- 95 E. Hylleras and B. Undheim; *Z. Physik* **65**, 759 (1930).
- 96 J. MacDonald; *Phys. Rev.* **43**, 830 (1933).
- 97 D. Craig; *Proc. Roy. Soc. (London)* **A200**, 474 (1950).
- 98 S. Boys; *Proc. Roy. Soc. (London)* **A201**, 125 (1950).
- 99 R. Nesbet; *Proc. Roy. Soc. (London)* **A230**, 312 (1955).
- 100 E. Davidson; *J. Chem. Phys.* **33**, 1577 (1960).
- 101 T.-S. Lee, D.M. York, and W. Yang; *J. Chem. Phys.*, **105**, 2744 (1996).
- 102 W. Yang; *Phys. Rev. Lett.*, **66**, 3547 (1992).
- 103 W. Yang; *Phys. Rev. A*, **44**, 7823 (1991).
- 104 R.G. Parr; *J. Chem. Phys.* **20**, 239 (1952).
- 105 R. Pariser and R. Parr; *J. Chem. Phys.* **21**, 466 (1953).

- 106 R. Pariser and R. Parr; *J. Chem. Phys.* **21**, 767 (1953).
- 107 R. Parr; *Quantum Theory of Molecular Electronic Structure*, Benjamin, New York, 1963.
- 108 J.A. Pople and D.L. Beveridge; *Approximate Molecular Orbital Theory*, McGraw-Hill, New York, 1970.
- 109 J.A. Pople and G.A. Segal; *J. Chem. Phys.* **43**, S136 (1965).
- 110 J.A. Pople and G.A. Segal; *J. Chem. Phys.* **44**, 3289 (1966).
- 111 D.P. Santry and G.A. Segal; *J. Chem. Phys.* **44**, 3289 (1966).
- 112 M.C. Bohm and R. Gleiter; *Theoret. Chim. Acta (Berlin)*, **59**, 127 (1981).
- 113 J.A. Pople, D.L. Beveridge, P.A. Dobosh; *J. Chem. Phys.* **47**, 2026, (1967).
- 114 J.E. Ridley and M.C. Zerner; *Theor. Chim. Acta (Berl.)*, **32**, 111 (1973).
- 115 J.A. Pople, D.P. Santry, and G.A. Segal; *J. Chem. Phys.* **43**, S129 (1965).
- 116 R. Sustmann, J.E. Williams, M.J.S. Dewar, L.C. Allen, P. von R. Schleyer; *J. Am. Chem. Soc.* **91**, 5350 (1969).
- 117 M.J.S. Dewar and W. Thiel; *J. Am. Chem. Soc.* **99**, 4899 (1977).
- 118 M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, and J.P. Stewart; *J. Am. Chem. Soc.* **107**, 3902 (1985).
- 119 R.V. Stanton, D.S. Hartsough, and K.M. Merz Jr; *J. Chem. Phys.*, **97**, 11868 (1993).
- 120 M.J. Field, P.A. Bash, and M. Karplus; *J. Comput. Chem.* **6**, 700 (1990).
- 121 U.C. Singh and P.A. Kollman; *J. Comput. Chem.* **7**, 718 (1986).
- 122 J. Gao; *J. Phys. Chem.* **96**, 537 (1992).
- 123 M.A. Thoompson; *J. Phys. Chem.* **99**, 4794 (1995).
- 124 F. Maseras and K. Morokuma; *J. Comput. Chem.* **16**, 1170 (1995).

- 125 *Reviews in Computation Chemistry Vol. 1* ed. K.B. Lipkowitz and D.B. Boyd, VCH, New York 1990.
- 126 E.B. Wilson Jr, J.C. Deciu, and P.C. Cross; *Molecular Vibrations*, McGraw-Hill, New York, 1955.
- 127 P. Scharfenberg; *Theor. Chim. Acta* **53**, 279 (1979).
- 128 R. Fletcher and C.M. Reeves; *J. Comput.* **1**, 149 (1964).
- 129 J.D. Head and M.C. Zerner; *Advances in Quantum Chemistry*, Vol. 20 Academic Press, New York, 1989.
- 130 C.G. Broyden; *J. Inst. Maths Applns.* **6**, 163 (1970).
- 130 R. Fletcher; *J. Comput.* **13**, 317 (1970).
- 131 D. Godfarb; *Maths. Comput.* **24**, 23 (1970)
- 132 D.F. Shanno; *Maths. Comput.* **24**, 647 (1970).
- 133 B.A. Murtagh and R.W.H. Sargent; *J. Comput* **13**, 185 (1970).
- 134 W.C. Davidon; *AEC Res. and Dev. Report ANL-5990* (1959).
- 135 R. Fletcher and M.J.D. Powell; *J. Comput.* **6**, 163 (1963).
- 136 P. Jorgensen and J. Simons; *NATO ASI series C* vol. 166 Reidel, Boston (1985).
- 137 R. Fletcher; *Practical Method of Optimization*, Vol. 1, Wiley, New York, 1980.
- 138 M.J.D. Powell; *Optimization in Action*, ed. L.C.W. Dixon, Academic Press, New York, 1976.
- 139 J.M. Haile; *Molecular Dynamics Simulations: Elementary Methods*, Wiley, New York, 1992
- 140 J. McCammon and Harvey; *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, 1987.
- 141 L. Verlet; *Phys. Rev.* **159**, 98 (1967).
- 142 L. Verlet; *Phys. Rev.* **165**, 201 (1968).
- 143 D. Beeman; *J. Comput. Phys.*, **20**, 130 (1976).

- 144 H. Berendsen, J. Postma, W. van Gunsteren, A. DiNola, and J. Haak; *Phys. Rev. A*, **81**, 3684 (1984).
- 145 H. Andersen; *J. Chem. Phys.*, **72**, 2384 (1980).
- 146 J. Ryckaert and G. Ciccotti; *J. Chem. Phys.*, **78**, 7368 (1983).
- 147 T. Schneider and E. Stoll; *Phys. Rev. B*, **13**, 1216 (1976).
- 148 W. Hoover, D. Evans, R. Hickman, A. Ladd, W. Ashurst, and B. Moran; *Phys. Rev. A*, **22**, 1690 (1980).
- 149 W.F. van Gunsteren and H.J.C. Berendsen; *Mol. Phys.*, **45**, 637 (1982).
- 150 E. Dickinson, *Chem. Soc. Rev.*, **14**, 421 (1985).
- 151 D.L. Ermak and J.A. McCammon; *J. Chem. Phys.*, **69**, 1352 (1985).
- 152 A. Iniesta and J.G. de la Torre; *J. Chem. Phys.*, **92**, 2015 (1990).
- 153 N. Metropolis and S. Ulam; *J. Am. Stat. Ass.*, **44**, 335 (1949).
- 154 Kelvin; *Phil. Mag.*, **2**, 1 (1901).
- 155 N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller; *J. Chem. Phys.*, **21**, 1087 (1953).
- 156 M.P. Allen and D.J. Tildsley; *Computer Simulation of Liquids*, Oxford University Press, Oxford, 1987.
- 157 K. L. Chung; *Markov Chains with Stationary State Probabilities*, Vol. 1, Springer, Heidelberg, 1960.
- 158 W.W. Wood; *Physics of Simple Liquids*, ed. H.N.V. Temperley, J.S. Rowlinson, and G.S. Rushbrooke, North Holland, Amsterdam, 1968.
- 159 C.A. Angell, J.H.R. Clarke, and L.V. Woodcock; *Adv. Chem. Phys.*, **48**, 397 (1981).
- 160 C.L. Brooks, M. Karlpus, and B.M. Pettitt; *Advances in Chemical Physics*, Vol. LXXI, Wiley, New York, 1988.
- 161 H. Dufnr, S.M. Kast, J. Brickmann, and M. Schlenkrich; *J. Comput. Chem.*, **18**, 660 (1997).
- 162 D.M. York, T.A. Darden, and L.G. Pedersen; *J. Chem. Phys.*, **99**, 8345 (1993).

- 163 T.E. Cheatham III, J.L. Miller, T. Fox, T.A. Darden, and P.A. Kollman; *J. Am. Chem. Soc.*, **117**, 4193 (1995).
- 164 E. Madelung; *Phys. Z.*, **19**, 524 (1918).
- 165 P. Ewald; *Ann. Phys.*, **64**, 253 (1921).
- 166 H. Lee, T. Darden, L. Pedersen; *Chem. Phys. Lett.*, **243**, 229 (1995).
- 167 D.M. York, A. Wlodawer, L.G. Pedersen, and T.A. Darden; *Biophysics*, **91**, 8715 (1994).
- 168 D.M. York, W. Yang, H. Lee, T. Darden, and L.G. Pedersen; *J. Am. Chem. Soc.*, **117**, 5001 (1995).
- 169 H. Lee, T.A. Darden, L.G. Pedersen; *J. Chem. Phys.*, **102**, 3830 (1995).
- 170 A.J. Stone; *Chem. Phys. Lett.*, **83**, 233 (1981).
- 171 M.O. Fenley, W.K. Olson, K. Chua, and A.H. Boschitsch; *J. Comput. Chem.*, **17**, 976 (1996).
- 172 J. Carrier, L. Greengard, and V. Rokhlin; *J. Sci. Stat. Comput.*, **9**, 669 (1988).
- 173 M. Born; *Z. Phys.*, **1**, 45 (1920).
- 174 J.G. Kirkwood; *J. Chem. Phys.*, **2**, 351 (1934).
- 175 L. Onsager; *J. Am. Chem. Soc.*, **58**, 1486 (1936).
- 176 F.T. Marchese, P.K. Mehrotra, and D.L. Beveridge; *J. Phys. Chem.*, **85**, 1 (1981).
- 177 H. Tanaka, K. Nakanishi, and N. Watanabe; *J. Chem. Phys.*, **78**, 2626 (1983).
- 178 W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein; *J. Chem. Phys.* **79**, 926 (1983).
- 179 M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, and J.P. Stewart; *J. Am. Chem. Soc.* **107**, 3902 (1985).
- 180 J.P. Stewart; *J. Comput. Chem.* **10**, 3902 (1985).
- 181 M.J. Frisch, J.E. DelBene, J.S. Binkley, and H.F. Schaefer III; *J. Chem. Phys.* **84**, 2279 (1986).

- 182 M. Mezei and J.J. Dannenberg; *J. Phys. Chem.* **92**, 5860 (1988).
- 183 C.J. Marsden, B.J. Smith, J.A. Pople, H.F. Schaefer and L. Random; *J. Chem. Phys.* **95**, 1825 (1991).
- 184 D. Feller; *J. Chem. Phys.* **96**, 6104 (1992).
- 185 M.W. Feyereisen, D. Feller, and D.A. Dixon; *J. Phys. Chem.* **100**, 2993 (1996).
- 186 H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, and J. Hermans; in *Intermolecular Forces*; B. Pullman Ed.; Reidel: Dordrecht Holland, 1981; p331.
- 187 J. Spooner and P. Hobza; *J. Phys. Chem.* **98**, 3161 (1994).
- 188 T.R. Dyke, K.M. Mack, and J.S. Muentert; *J. Chem. Phys.*, **66**, 498 (1977).
- 189 L.A. Curtiss, D.J. Frurip, and M. Blander; *J. Chem. Phys.*, **71**, 2703 (1979).
- 190 M. Watanabe and M. Karplus; *J. Chem. Phys.* **99**, 8063 (1993).
- 191 P.G. Jansen and W.J. Stevens; *J. Chem. Phys.* **84**, 3271 (1986).
- 192 J.C. Contador, M.L. Sanchez, M.A. Aguilar, and F.J.O. del Valle; *J. Chem. Phys.* **104**, 5539 (1996).
- 193 M.W. Jurema and C. Shields; *J. Comput. Chem.*, **14**, 89 (1993).
- 194 R.S. Becker and G. Kogan; *Photochem. Photophys.* **31**, 5 (1980)
- 195 I. Baraldi, M.C. Bruni, M.P. Costi, and P. Pecorari; *Photochem. Photophys.* **52**, 361 (1990).
- 196 A. Broo; *Chem. Phys.* **174**, 127 (1993).
- 197 P.R. Callis; *Annu. Rev. Phys. Chem.* **34**, 329 (1983).
- 198 P.R. Callis; *Photochem. Photobiol.* **44**, 315 (1986).
- 199 A. Broo and Holemen, Submitted
- 200 J. Lorentzon, M.P. Fölscher, and B.O. Roos; *J. Am. Chem. Soc.* **117**, 9265 (1985).
- 201 C.A. Sprecher and W.C. Johnson; *Biopolymers* **16**, 2243 (1977).

- 202 M. Karelson and M.C. Zerner; *J. Chem. Phys.* **96**, 6949 (1992).
- 203 M.T. Fujii, T. Tamura, and N. Mikami; *Chem. Phys. Lett.* **126**, 583 (1986).
- 204 I.R. Gould, N.A. Burton, R.J. Hall, and I.H. Hillier; *J. Mol. Struct. (THEOCHEM)* **331**, 147 (1995).
- 205 D.A. Estrin, L. Paglieri, and G. Corongiu; *J. Phys. Chem.* **98**, 5653 (1994).
- 206 S.A. Williams, C.N. Renn, and P.R. Callis; *J. Phys. Chem.* **91**, 2730 (1987).
- 207 W.D. Edwards and M.C. Zerner; *Theor. Chim. Acta* **72**, 347 (1987).
- 208 C.C.J. Roothaan; *Rev. Mod. Phys.* **32**, 179 (1960).
- 209 J.E. Ridley and M.C. Zerner; *Theor. Chim. Acta* **42**, 223 (1976).
- 210 E.C. Lim; *J. Phys. Chem.* **90**, 6770 (1986).
- 211 M.M. Karelson, T. Tamm, A.R. Katritzky, S.J. Cato, and M.C. Zerner; *Tetrahedron Comput. Method* **2**, 295, (1989).
- 212 A. Klamt and G. Schüürmann; *J. Chem. Soc., Perkin Trans.* **2**, 799, (1993).
- 213 C.J. Cramer and D.G. Truhlar; *J. Am. Chem. Soc.* **113**, 8305, (1991).
- 214 J. Tomasi and M. Persico; *Chem. Rev.* **94**, 2027, (1994).
- 215 K.K. Stavrev, M.C. Zerner, and T. Meyer; *J. Am. Chem. Soc.* **117**, 8684, (1995).
- 216 P. Ford, De F. P. Rudd, R. Gaunder, and H. Taube; *J. Am. Chem. Soc.* **90**, 1187, (1968).
- 217 J.C. Curtis and T. Meyer; *J. Inorg. Chem.* **21**, 1562, (1982).
- 218 H.C. Stynes and A. Ibers; *J. Inorg. Chem.* **10**, 2304 (1971).
- 219 D.P. Rillema, D.S. Jones, and H.A. Levy; *J. Chem. Soc., Chem. Commun.* 849, (1979).
- 220 T. Meyer; *J. Pure Appl. Chem.* **58**, 1193, (1986).
- 221 S. Woitellier and J.P. Launay; *Joachim, C, Phys.* **131**, 481, (1989).

- 222 A. Jurus, V. Balzani, F. Barigelleti, S. Campagna, P. Belser, and A. von Zelewsky; *Coord. Chem. Rev.* **84**, 85, (1988).
- 223 M.J. Ondrechen, M.A. Ratner, and D.E. Ellis; *J. Am. Chem. Soc.* **103**, 1656, (1981).
- 224 D.E. Richardson and H. Traube; *J. Am. Chem. Soc.* **105**, 40, (1983).
- 225 M.M. Karelson and M.C. Zerner; *J. Phys. Chem.* **96**, 6949, (1992).
- 226 A. Broo; *Chem. Phys.* **174**, 127, (1993).
- 227 A. Broo and S. Larsson; *Chem. Phys.* **161**, 363, (1992).
- 228 M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T.A. Keith, G.A. Peterson, J.A. Montgomery, K. Raghavachari, M.A. Allaham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Detrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez, and J. Pople; *Gaussian* (Gaussian Inc., Pittsburgh, PA).
- 229 A. Warshel and M. Levitt; *J. Mol. Biol.* **103**, 227, (1976).
- 230 U.C. Singh and P.A. Kollman; *J. Comput. Chem.* **7**, 718, (1986).
- 231 M.J. Field, P.A. Bash, and M. Karplus; *J. Comput. Chem.* **10**, 700, (1990).
- 232 F. Maseras and K. Morokuma; *J. Comput. Chem.* **16**, 1170, (1995).
- 233 A. Broo, G.M. Pearl, and M.C. Zerner; *J. Phys. Chem.* **101**, 2478, (1997).
- 234 R.G. Parr; "Quantum Theory of Molecular Electronic Structure", (1963) Benjamin, Boston.
- 235 M.C. Zerner; "In Problem Solving in Computational Molecular Sciences", (1997), NATO ASI Series, Dordrecht.
- 236 J.E. Ridley and M.C. Zerner; *Theoret. Chim. Acta*, **32**, 111, (1973).
- 237 M.C. Zerner, G.H. Loew, R.F. Kirchner, U.T. Mueller-Westerhoff; *J. Am. Chem. Soc.*, **102**, 589, (1980).

- 238 J.C. Culberson, P. Knappe, N. Rosch and M.C. Zerner; *Theoret. Chim. Acta*, **71**, 21, (1987).
- 239 J.D. Baker and M.C. Zerner; *Chem. Phys. Lett.*, **175**, 192, (1990).
- 240 H. Nakatsuji, M. Komari, O. Kitao; *Chem. Phys. Lett.*, **142**, 447, (1987).
- 241 H. Nakatsuji, O. Kitao and T. Yonezawa; *J. Chem. Phys.*, **83**, 723, (1985).
- 242 H. Nakatsuji, J.-Y. Hasegawa and M. Hada; *J. Chem. Phys.*, **104**, 2321, (1996).
- 243 J.F. Stanton and R.J. Bartlett; *J. Chem. Phys.*, **101**, 7029, (1993).
- 244 for a review of couple-cluster methods, see R.J. Bartlett and J.F. Stanton 1994, in *Reviews of Computational Chemistry*. Vol 5 ed. K. Lipkowitz and D.B. Boyd, VCH, New York.
- 245 P.E. Siegbahn, A. Heiberg, B.O. Roos, and B. Levy; *Phys. Scr.*, **21**, 323, (1980).
- 246 B.O. Roos, P.R. Taylor and P.E. Siegbahn; *Chem. Phys.*, **48**, 157, (1980).
- 247 J. Karwowski; *Intern J. Quantum Chem.*, **51**, 425, (1994).
- 248 D. Bielinska-Waz and J. Karowski; *Phys. Rev. A*, **55**, 1067, (1995).
- 249 D. Bielinski-Waz and J. Karowski; *Advances in Quantum Chemistry*, Vol 28, pp 168, (1997) Academic Press, San Diego.
- 250 P.J. Campagnola, D.J. Lavrich, M.J. DeLuca and M.A. Johnson; *J. Chem. Phys.*, **94**, 5240, (1991).
- 251 A. Baczynski and D. Radomska; *Journal of Fluorescence*, **5**, 91, (1995).
- 252 M. Klessinger and J. Michl; *Lichtabsorption and Photochemie Organischer Molekule*, VCH, 1989, New York
- 253 D.A. McQuarrie; *Statistical Mechanics*, Harper Collins, 1976, New York.
- 254 J. Blair, K. Krough-Jespersen and R. Levy; *J. Am. Chem. Soc.*, **111**, 6948, (1989).

- 255 R. Levy, D.B. Kitchen, J. Blair, K. Krough-Jespersen; *J. Phys. Chem.*, **94**, 4470, (1990).
- 256 J. Gao; *J. Am. Chem. Soc.*, **116**, 9324, (1994).
- 257 K. Coutinho, M.J. de Oliveira, and S. Canuto; *Intern. J. Quantum Chem.*, **66**, 249 (1998).
- 258 V. Luzhkov and A. Warshel; *J. Am. Chem. Soc.*, **112**, 4491, (1991).
- 259 A. Warshel; *J. Phys. Chem.*, **83**, 4912, (1979).
- 260 S.E. DeBolt and P.A. Kollman; *J. Am. Chem. Soc.*, **112**, 7515, (1990).
- 261 J. Zeng, N.S. Hush and J.R. Reimer; *J. Phys. Chem.*, **100**, 9561, (1996).
- 262 H. Eyring, J. Walter and G.E. Kimball; *Quantum Chemistry*, 1944, John Wiley and Sons.
- 263 S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta and P. Weiner; *J. Am. Chem. Soc.*, **106**, 765, (1984).
- 264 K.B. Wiberg; *Tetrahedron*, **24**, 1083, (1968).
- 265 M.C. Zerner and R.G. Parr; *J. Chem. Phys.*, **69**, 3858, (1978).
- 266 K. Jug; *J. Comp. Chem.*, **5**, 555, (1984).
- 267 J.B. Foresman, M. Head-Gorden, J.A. Pople, M.J. Frish; *J. Chem. Phys.*, **96**, 135, (1992).
- 268 *UV atlas of organic compounds*; Organischer Verbindungen, Plenum Press, New York, (1966-).
- 269 Anders Broo, Greg Pearl and Michael C. Zerner; *J. Phys. Chem.*, **101**, 2478, (1997).
- 270 N. Mataga and K. Nishimoto; *Z. Chem. Phys.*, **13**, 140, (1957).
- 271 M.M. Karelson and M.C. Zerner; *J. Chem. Phys.*, **96**, 6949, (1992).
- 272 N. Rösch and M.C. Zerner; *J. Phys. Chem.*, **98**, 5817, (1994).
- 273 M.M. Karelson and M.C. Zerner; *J. Am. Chem. Soc.*, **112**, 9405, (1990).

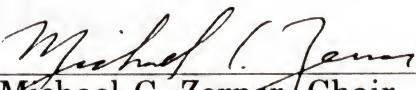
BIOGRAPHICAL SKETCH

I was born on October 15, 1971 and grew up in Corydon a small town in Southern Indiana, approximately 12 miles west of Louisville, Kentucky. I attended the Corydon Central High School, my graduating class had approximately 120 people.


I attended Indiana University-Purdue University at Indianapolis (IUPUI), where I worked with Dr. Kenny Lipkowitz for three years on several different computational chemistry problems. I graduated in the summer of 1997 with an ACS certified BS degree in Chemistry with a minor in Mathematics and Physics. While attending IUPUI I was also a chemistry laboratory and recitation instructor for two and a half years.

I have been working with Dr. Zerner here at the University of Florida for the last four years. During my stay here I visited Göteborg Sweden for 6 weeks during the summer of 1996 to work with Dr. Anders Broo. I have also been maintaining the ZINDO program package for the last two years and consulting Molecular Simulations Inc. with developing a Graphical User interface for ZINDO along with technical support. I have also been assisting Dr. Erik Deumens as a system administrator for the QTP computing environment for the last three and a half years.

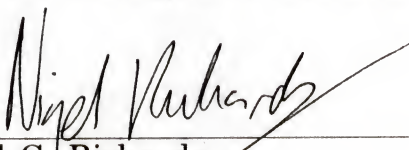
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Michael C. Zerner, Chair
Professor of Chemistry


I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Yngve N. Öhrn
Professor of Chemistry

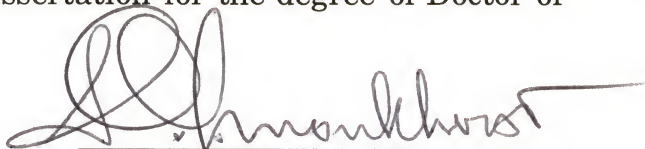
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Nigel G. Richards
Associate Professor of
Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


John R. Sabin
Professor of Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

A handwritten signature in dark ink, appearing to read 'H. Monkhorst', written over a horizontal line.

Hendrik J. Monkhorst
Professor of Physics

This thesis was submitted to the Graduate Faculty of the Department of Chemistry in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 1998

Dean, Graduate School